Aalto University
School of Economics

Models of network formation - Implications of network approach in labor markets

Economics

Master's thesis

Elisa Tiilimäki

2010

Department of Economics
Aalto University
School of Economics

Aalto University
School of Economics

MODELS OF NETWORK FORMATION

Implications of network approach in labor markets

Master's Thesis
Elisa Tiilimäki
26.04.2010
Economics

Approved by the Head of the Economics Department __.__.2010 and awarded the grade

____

1. examiner 2. examiner

AALTO UNIVERSITY SCHOOL OF ECONOMICS
Department of Economics
Elisa Tiilimäki

ABSTRACT
April 24 2010

THE MODELS OF NETWORK FORMATION
Implications of network approach in labor markets

Social networks play an important role in many significant economic relationships, and further in a wide range of interactions, including the transmission of job information, how much education one gets and how diseases spread. Networks affect the flow of information, which makes it critical to understand which networks structures are likely to emerge in an economy and how social network structures affect behavior. Networks have been long present in the sociology but the importance of networks in economics has not been acknowledged until the early 1990's. Furthermore, the amount of social networking sites has increased rapidly during the past few years, which makes this field of research a very current issue.

In this literature review, I dedicate a substantial part to study the various models of network formation. I analyze the differences between the distinct groups of network formation. My other aim is to examine the very effects of networks, i.e. the emerged network structures to the economy. Here, I address the study especially on how networks affect job finding and how one can explain duration dependence of unemployment and persistent inequality of wages with the presence of networks. I compare the search theory with the labor market theories assuming network approach. My goal is thus to show the importance of the network approach in economic theories with the labor market example.

MAIN FINDINGS
The first and the most known model in random graph literature is the static Erdös and Rényi (1959) model where the probability of link formation is Poisson distributed. Basic random graph model fails to replicate many features present in the real social networks such as clustering and "small-world" effects. One major shortcoming is the assumption of independent link formation; therefore e.g. models of homophilous networks have been developed. Random-graph models do not analyze the link formation process on agent level or the stability and efficiency as does strategic models; consequently, strategic models are more prevalent in the economic literature. Jackson and Wolinsky (1996) introduce the equilibrium concept of pairwise stability in their strategic formation model. Another pioneering one is the game theoretic, noncooperative model by Bala and Goyal (2000). Applications of game-theoretic models include dynamic models, network formation with heterogeneous agents and with endogenous link strength.

Networks are a significant means of acquiring a job. There are labor market models implementing network approach with both fixed and variable network structure. Network approach explains the duration dependence of unemployment by the fact that longer history of unemployment is more likely to come when the direct and indirect connections of an agent are unemployed. Inequality explicated by networks appears in the difference of wage levels and dropout rates. Network approach provides a complementary theory to search model in deliberating labor market phenomena. The key finding and the key merit of the network approach is the fact that it does not assume anonymous markets; the network structure determines which agents meet and operate together in the market. The impact of networks on information transmission patterns is supposedly significant because access to information is heavily influenced by social structure. Networks most likely reduce the problem of information asymmetry but it also causes individuals to have asymmetric positions in the economy with respect to access to information, which might result in persistent inequality between individuals. Networks might in fact be one source of market frictions. Within this light, networks deserve more attention in economic theories. One should also account for the growing role of Internet and institutions in the research.

Key words: Networks, network formation model, job search, labor market theories, anonymous markets, information transmission

TIIVISTELMÄ
24.4.2010

Sosiaalisilla verkostoilla on merkittävä osa tärkeissä taloudellisissa suhteissa sekä monenlaisessa vuorovaikutuksessa, kuten työpaikoista kertovan tiedon välittymisessä. Verkostot vaikuttavat myös yksilöiden koulutustasoon, tautien leviämiseen ja siihen, mitä kieliä puhutaan. Koska verkostot vaikuttavat tiedon kulkuun, on tärkeää ymmärtää, minkälaisia verkostorakenteita todennäköisesti syntyy ja miten verkosto–rakenne vaikuttaa taloudellisten toimijoiden käyttäytymiseen. Verkostoja on jo kauan tutkittu sosiolo–giassa, mutta verkostojen merkitys taloustieteessä on tunnustettu vasta 1990-luvun alussa. Lisäksi sosiaalisten verkostoitumissivustojen määrä on viime vuosina kasvanut voimakkaasti, minkä vuoksi aihe on erittäin ajankohtainen.

Käsittelen tässä kirjallisuuskatsauksessa laajasti erilaisia verkostonmuodostumismalleja ja pyrin tarkastelemaan eri muodostumismalliryhmien välisiä eroja. Edelleen tavoitteenani on tutkia verkostojen ja verkostorakenteiden vaikutusta talouteen. Keskityn tutkimaan sitä, miten verkostot vaikuttavat työn etsimiseen ja työllistymiseen, ja sitä, miten verkostonäkökulmalla voidaan selittää työttömyyden kestoa ja myös palkkojen epäsuhtaisuutta eri ryhmien välillä. Vertaan etsintäteoriaa (*search theory*) ja verkostonäkökulmaan perustuvia työmarkkinamalleja. Pyrin näin osoittamaan verkostonäkökulman tärkeän merkityksen taloudellisissa teorioissa.

TULOKSET
Tunnetuin satunnaiseen muodostumiseen perustuva verkostomalli (*random graph*) on Erdös ja Rényin (1959) malli, jossa linkin muodostumistodennäköisyys noudattaa Poisson-jakaumaa. Random graph -mallit eivät pysty kuvaamaan oikeiden verkostojen ominaisuuksia, kuten klusteroitumista ja *small-world*-ilmiötä. Suuri puute on myös se, että monissa malleissa oletetaan linkkien muodostuvan itsenäisesti. Poikkeuksia ovat mm. homofiiliset verkostonmuodostumismallit, sekä eksponentiaaliset *random graph* –muodostumismallit. Random graph mallit tarkastelevat linkinmuodostumisprosessia verkostotasolla kun taas strategiset mallit kuvaavat prosessia agenttitasolla; tärkeänä osana on syntyneen verkoston tasapainon ja tehokkuuden tutkiminen. Jacksonin ja Wolinskyn (1996) strategisen muodostumisen malli perustuu parilliseen tasapainoon (*pairwise stability*). Toinen uraauurtava malli on Bala ja Goyalin (2000) peliteoreettinen malli. Peliteoreettisia muunnelmia ovat mm. dynaaminen malli, heterogeenisten agenttien malli sekä malli, jossa linkin vahvuus on endogeeninen muuttuja.

Verkostot ovat merkittävä tekijä työllistymisessä. Verkostonäkökulmaan perustuvat työmarkkinamallit voidaan jakaa kahteen ryhmään; kiinteän verkostorakenteen olettavat sekä vaihtuvan verkostorakenteen olettavat mallit. Verkostoteoria selittää työttömyyden kestoriippuvuutta (*duration dependence*) sillä, että pidempiaikainen työttömyys on todennäköisempää silloin kun agentin suorat ja epäsuorat kontaktit ovat työttömiä. Palkkatasoerot eri ryhmien, välillä johtunevat ryhmien eroista työmarkkinoiden ulkopuolelle jättäytymisessä.

Verkostonäkökulma tarjoaa etsintäteorian ohella täydentävän teorian työmarkkinailmiöiden selittämiseen. Verkostoteorian yksi pääansioista on kritiikki anonyymejä markkinoita kohtaan; verkostorakenne määrittelee ketkä tapaavat ja toimivat riippuvuussuhteessa markkinoilla. Verkostojen vaikutus tiedonkulkuun on tiettävästi merkittävä, sillä sosiaalinen rakenne määrittää tiedon saatavuutta. Verkostot myös vähentänevät informaatioasymmetriaa, mutta toisaalta verkostojen takia yksilöiden informaatiotaso taloudessa on asymmetrinen. Tämä voi johtaa pysyvään epätasa-arvoon taloudellisten toimijoiden välillä, verkostot voivat siis aiheuttaa merkittävästi markkinakitkaa. Verkostot ansaitsevat siis enemmän huomiota taloudellisissa teorioissa. Myös Internetin ja muiden instituutioiden kasvava rooli tiedon välittäjänä ja verkostorakenteiden muokkaajana tulisi huomioida laajemmin.

Avainsanat: Verkostot, verkostomuodostumismalli, työmarkkinateoriat, anonyymit markkinat, informaation välittyminen

Contents

1. Introduction 3
    1.1. The impact of networks to the economic outcome 3
    1.2. Research question 5
        1.2.1. Networks in the literature 6
    1.3. Results 7
2. Networks – an introduction 9
    2.1 Overview of the literature and terminology 10
    2.2. Modeling networks 11
3. Nonstrategic models of network formation 15
    3.1. Random-Graph Models of Network Formation 16
        3.1.1. Basic Random-Graph Models 16
        3.1.2. Properties of random-graph networks and real social networks compared 17
    3.2. Extensions of the basic random-graph models 18
        3.2.1. A small-world model 19
        3.2.2. Fixed degree distributions 20
        3.2.3. Preferential attachment 21
        3.2.4. Homophilous networks 21
        3.2.5. Exponential random graph models 24
    3.3. Conclusion 26
4. Strategic and game theoretic modeling of network formation 29
    4.1. Comparison of the random-graph and strategic networks 29
        4.1.1. Overview of the strategic form models 30
    4.2. Strategic network formation model 31
        4.2.1. The Connections Model 32
        4.2.2. The Co-Author Model 34
        4.2.3. Stability and efficiency 34
    4.3. Game theoretic network formation models 35
        4.3.1. One and two-sided link formation 35
    4.4. Discussion about the different equilibrium concepts 39
5. Other game theoretic models of network formation 42
    5.1. Dynamic models 42
    5.2. Endogenous link strength and heterogeneity across agents 43
        5.2.1. Link strength as an endogenous variable 44
        5.2.2. Heterogeneous agents – asking and giving model 45
    5.3. Discussion 47

6. Networks and labour markets 48
6.1. Characteristics of networked labor markets 48
6.2. Labor market models 50
6.2.1. Introduction to search and matching models 50
6.2.2. The network approach 51
6.3. Network approach explaining phenomena in labor markets – comparison with search approach 55
6.3.1. Duration dependence of unemployment 55
6.3.2. Inequality in terms of wage difference and dropout levels 59
6.4. Discussion 62
7. Conclusion 64
7.1. The models of network formation compared 64
7.2. Networks and labor markets 65
7.3. Networks and flow of information 67
7.4. Suggestions for future research 68
References 70
Appendix 74

List of pictures

Picture 1. Undirected and directed network 12
Picture 2. The degree of a graph. 13
Picture 3 Triadic closure 14
Picture 4. Erdös-Renyi Random Graph G 16
Picture 5. Random rewiring procedure 19
Picture 6. Illustration of homophily 22
Picture 7. Strict Nash networks, one-way model 38

List of tables

Table 1. Measured and random graph clustering coefficients. 17
Table 2. Networks as a source of hire. 49

1. Introduction

Research on network formation within economics is generally motivated by the observation that network structures play an important role in the organization of some significant economic relationships. The wide range of networked interactions include the buying and selling of many goods and services, the transmission of job information and informal insurance networks (Jackson and Wolinsky 1996). Social networks are also important in determining how diseases spread, which languages are spoken, how much education one gets and eventually, the likelihood of succeeding professionally. (Matthew Jackson 2008) Theoretical models of network formation have also high-lighted the role of networks in explaining phenomena such as stock market volatility, collective action, the career profiles of managers, and the diffusion of new products, technologies and conventions (Bala and Goyal 2000). Obviously, the influence mechanism of the networks on the economic relationships is the fact that networks affect the flow of information. This makes it critical to understand which networks structures are likely to emerge in an economy and how social network structures affect behavior.

Networks have been studied in the academic literature for many decades, most notably in sociology. For a long time, economists did not pay attention to the importance of networks in explaining economic phenomena. Around the turn of the century, it became clear that that neoclassical economic theory could not explain some patterns in labor markets and in the exchange of goods. From this pont onwards, the number of economic models with network features has exploded.

1.1. The impact of networks to the economic outcome

If the pattern of trade derives from actions taken in the manner expected in the neoclassical theory where economic agents are expected constantly to be initiating and ceasing interaction with one another as they search for the best deal, the network structures would be seemingly irrelevant when considering the economic outcome (Zuckerman 2003). However, social networks may have a significant impact on the economic outcome, and the economic process itself modifies the networks (Kirman 1997). How and why networks matter is nonetheless subject of significant debate. Granovetter (2005), a sociologist, suggests that there are three main reasons why networks affect the economic outcome. First, social networks affect the flow and the quality of information since actors do not believe impersonal sources and instead rely on people they know. Second, social networks are an important source for reward and for punishment. Third, trust emerges. By trust Granovetter means the confidence that other will do the “right” thing despite a clear balance of incentives to the contrary. The first aspect is the most appealing and relevant from the economic point of view and thus will be the one that I pay attention to in this work – although the emergence of trust might have significant implications for the game theoretic modeling. Networks as a source of information can in fact reduce the problem of information asymmetry and the one of moral hazard. One interesting example of networks' impact on market outcome is by Kirman and Vignes (1991). They find out that in Marseille Fish market, high proportion of buyers refrains from searching beyond their usual source of fish and even pay a higher price than in equilibrium as a result. This commodity market may be described as a patterned network of exchange. Kirman and Vignes find that in this networked market there is significant price dispersion and individual demand curves are not downward sloping. In a networked market trading with the same partners reduces the moral hazard problem since the quality of the goods is probably known. Also, networks might lower the search costs, which allows agents to pay a higher price.

Furthermore, in most real labor markets social networks play a key role. Prospective employers and employees prefer to learn about one another from personal sources whose information they trust. It has obvious links to theories of asymmetric information, with the difference that there is what one might call bilateral asymmetry—both employer and employee have information about their own "quality" that the other needs (Granovetter 2005). Network perspective is able to explain further different fundamental features of labor markets such as the duration (or resistance) of unemployment known as duration dependence (Calvó-Armengol and Jackson 2004) which traditional search models may not fully clarify. Also, the network approach enables more realistic assumptions e.g. in that unemployment is not just due to frictions in the market[1].

Social networks are not restricted to physical networks. The popularity of social networking sites has grown rapidly during the past few years and the field of these online networks is evolving continuously. Facebook is perhaps the most popular social networking site at the moment, measured in terms of number of user accounts. Other popular sites include Twitter, LinkedIn and Youtube. The impact of these virtual networks on the economic outcome might be hard to decipher due to the various mechanisms the online information sharing and communication in networks affects the outcome. Another problem might be the difficulty to actually distinguish between the effects of physical and online networks. These kinds of social networks might offer new cost effective ways to communicate which can in turn lower searching costs and transaction costs and promote new ideas, that is, increase productivity and economic welfare. A recent example of the economic impacts of

[1] Neoclassical theory assumes that economy tends to full employment equilibrium, if the labor market worked properly. If there was unemployment in the economy, the wages would fall, which would result in increased demand for labor and finally, equilibrium is restored at full employment. Here, any unemployment is purely voluntary. See Picture 7 in appendix. (New Palgrave Dictionary of Economics)

social media tools is the study conducted by FinNode USA, Finpro and Tekes in cooperation with Collaborative Strategies consultant company and Elearning! Magazine (9/2009). According to the study, social medias such as Facebook and Twitter create opportunities by altering the business operations and workers' approach, which have not been exploited in a large scale at corporate level yet. One exception taking advantage of social network tools is a company named Harvard Medical Center that was able to reduce the number of denounced from 600 to 80 by utilizing of social medias in the idea creating process. Among the research also corporate managers begin to understand the impact of social collaboration as a means to make the business more effective.

1.2. Research question

I dedicate a substantial part of my thesis to a survey of models of network formation. My other aim is to study the effects of networks in general and within the particular case of labor markets. The emphasis in the second part is on how networks affect the probabilities of finding jobs and the overall labor market outcome. I discuss how different models of network formation model the transmission of information on vacancies between labor market participants. In my mind, the assumption of full information is the most problematic one in the competitive model typically used in analyzing labor markets. The role of networks in structuring and controlling information flows within these markets provides a key for understanding why some workers succeed and some others are left out in actual labor markets.

The full information results from the neoclassical assumption of anonymous markets where all the actors in the market meet but the identity of the actors is not relevant to the market outcome. Particularly, this assumption in the neoclassical model is disproved by network models. Since all actors are not interconnected, every agent's information and possibilities to economic transactions depend who they know and deal with. Indeed, critique of Zuckerman (2003) about neoclassical model assuming utility maximizing agents (agents searching for the best deal) is not essentially what network models are criticizing.

Another interesting subject of study is the influence of online social medias, including Facebook, to the economic outcome. However, being such a recent topic the literature available is supposedly limited and reliable results from the effect of online social medias are hard to obtain due to short time span. In addition, distinguishing between the effects of real and virtual networks might be troublesome. Albeit social medias are part of most of the modern world individuals' everyday lives and they affect the way in which people communicate, I believe that the influence of social networks on the labor market outcome is as good area of study as social medias. All the economic actors are participants in the networked labor markets whereas not all utilize networking sites. The findings of networks' effect on labor market outcome can probably be generalized to apply also in the context of virtual networks.

The network formation models can be divided into two main classes based on the formation approach adopted and further into several subclasses. I first speak about non-strategic, random models of network formation where the models apply also to the formation of nonsocial networks, such as technical networks. The majority of my work is devoted to strategic models of network formation where the objective is to understand the formation of networks in terms of individual incentives. On my mind, the quintessence of networks from the economic perspective is the significant role of social and economic networks in explaining the economic outcome as compared to the conventional assumption of anonymous markets; consequently, social relations and informal institutions occupy a center stage in a large body of economics research.

I consider both random and strategic form models because both approaches are present in the labor market models with network perspective. I have also included extensions of the game theoretic models that I find relevant within the labor market context but have not yet been implemented, most probably due to short lifespan of the field of research.

1.2.1. Networks in the literature

There is an extensive literature on social networks from a sociological perspective (Jackson and Wolinsky 1996). The economic literature about networks is not restricted to the study of the formation of networks, but there are various studies for instance about transportation and delivery networks, e.g. routing of airlines by Hendricks, Piccione and Tan (1995). Microeconomic theory has applied network structures for issues such as the internal organization of firms (e.g. Boorman 1975), employment search (Montgomery 1991, Calvó-Armengol and Zenou 2005), and information transmission (Goyal 1993). However, one of the main contributions of the economic literature within this field has still been the study of endogenous formation of social networks by self-interested economic agents, where agents are economic or social actors (Bloch and Jackson 2006). Also, there is a formal game theoretic literature with cooperative approach including e.g. games with communication structures (Aumann and Myerson 1988, Dutta et al. 1998). Networks that have been subject to research in the economic literature include collaboration graphs, who talks to whom-graphs, and information linkage graphs. There are also technological networks and networks in the natural world; the architecture of the technical networks, such as ICT networks, most probably affect to the way in which information flows but I will nonetheless not focus on the underlying factors affecting information transmission. A more detailed list about the literature addressing the formation of networks is found in chapter 2.

1.3. Results

I discuss the network formation models that are the most important and most prevalent in the literature bearing in mind the labor market perspective. The best known and the first model in random graph literature is the static Erdös and Rényi (1959) model where the probability of link formation is Poisson distributed. Basic random graph model fails to replicate many features present in the real social networks such as clustering and "small-world" effects that many extensions, the dynamic scale-free model by Barabàsi and Albert (1999) for instance, try to take account for. In addition, one major shortcoming of many random graph models is the fact that they assume independent link formation. To fill this gap, Golub and Jackson (2008) have developed a model of homophilous networks and e.g. Snijders et al. (2006) speak about exponential random graph models also assuming nonindependent formation of links. Nonetheless, economics seems to be more interested in studying why networks form; what are the incentives for individuals to form certain links and how does the formation of certain network affect the economic welfare. That is why strategic form models are more prevailing in the economic literature. However, random graph models suit better to illustrate certain link formation processes that are important from the economic point of view. Random models have their place describing phenomena such as the spreading of pandemic flu, the diffusion of new products and technologies. In random-graph models that I present in the third chapter the approach is quite mathematical and the focal point is the analysis of network structures. Random-graph models do not provide the means to properly analyze the actual link formation process or the stability and efficiency as do strategic models, but they should still be included in the study because the structural aspects do play a significant role in the transmission of information.

I consider two fundamental strategic form models. Jackson and Wolinsky (1996) present a strategic formation model where they introduce the equilibrium concept of pairwise stability. Another pioneering one is the game theoretic, noncooperative network formation model by Bala and Goyal (2000). Nash equilbrium in Bala and Goyal and in many other models is a quite hard concept to deal with compared to pairwise stability because many Nash equilibrium networks, such as empty networks are not efficient, thus not welfare increasing or rational with respect to individual incentives because forming at least one link is shown to be welfare increasing. Pairwise stability is an easy but a rather weak concept to employ although it may provide strong results by narrowing the set of graphs substantially what Nash equilibrium fails to do. Applications of game-theoretic models include dynamic model of Bala and Goyal (2000), network formation with heterogeneous agents (e.g. Galeotti and Goyal 2002) and link formation with endogenous link strength (Bloch and Dutta 2009).

The role of social networks in explaining various economic phenomena has become more notable recently. For example, networks are a significant means in job search: among Montgomery (1991) over 50 % of individuals have found their current job through network of contacts, although there are differences between races and gender. Now there is also a move towards network approach in labor market theories because more commonly used search theoretic models do not account for the importance of social networks in employment process. Therefore, they are unable to rationalize certain phenomena such as duration dependence of unemployment and persistent inequality prevalent in the markets. There are labor market models implementing network approach with both fixed and variable network structure. Simpler models of fixed network structure include Calvó-Armengol and Jackson (2004) and Montgomery (1991), variable network structure model is introduced by Calvó-Armengol (2004).

Calvó-Armengol and Jackson (2004) explain the expected probability of acquiring a job decreasing in the length of time that an agent has been unemployed (duration dependence) by the fact that longer history of unemployment is more likely to come when the direct and indirect connections of an agent are unemployed. Except in duration dependence of unemployment, inequality appears e.g. in the difference of wage levels and dropout rates. Calvó-Armengol and Jackson (2007) show that any path connected agents' wages are positively correlated and that agent with many links ought to receive multiple job offers compared to agent with only few links. Persistent inequality can arise between two similar initial conditions with just different initial employment conditions because the difference in employment conditions leads to drop-out rates of unequal height. Dropping out hurts the prospects of other agents in the network further, this results in deeper wage differential between the groups. Measuring the impact of using networks in job search to wage level is quite complicated, though. Recognizing the effect of networks to employment and wage levels can provide significant policy guidelines in decreasing the inequality in labor markets by affecting network formation by means of transfers, for example, to improve the position of isolated individuals with nonexistent contacts.

As Granovetter (1974) has suggested, the impact of networks in information transmission patterns is supposedly significant, and the same thing also applies to the labor markets; it is the diffusion pattern of job information that determines who will be employed and who will not. One could also argue that networks are not that relevant in causing certain phenomena. But it seems that with network approach, one is able to explain quite well many phenomena present in the economy. The theories of network formation and the labor market theories with network approach do not strive for providing universally applicable theories overriding previous theories such as search model in labor markets, but for offering complementary explanations to fill the shortcomings of some of assumptions of neoclassical theories. Thus search models or efficiency wage models should not be abandoned altogether but network approach just adds to basic theoretical models and helps to understand causalities in the markets because clearly, networked markets function differently than those where individuals are not dependent on the structure of the network of contacts but can act independently. This kind of market structure with no links but only independent actors is supposedly only to be seen in theories.

This paper is organized as follows; I introduce the basic concepts of networks in the second part and in part 3, I go through the random graph model of network formation and the most important applications. In chapter 4, I present the basic strategic and game theoretic models that are the most prevalent in literature. Chapter 5 is devoted to the extensions of the basic game theoretic models which I consider essential from the labor market perspective. In part 6 I turn to speaking about labor markets with network approach. I discuss some labor market models using network approach and think about how one can explain some phenomena in labor markets with the existence of networks. Chapter 7 concludes.

2. Networks – an introduction

The network literature within the field of economics dates back to the 1950's; both the focus of the literature and the terminology within the analysis of formation of networks have evolved during the past decades. I will go through the different terms present in the literature so to elucidate the subject at hand. I briefly present the key characteristics of coalitions, the predecessor of networks as well.

Most models view networks as either non-directed or directed graphs[2]; which type of graph is more appropriate depends on the context. Many social and economic relationships are reciprocal or require the consent of both parties to form; there are also enough applications especially in the context of social networking sites that take a directed form so that presenting both kinds of networks

[2] In mathematical graph theory, a graph is a finite set of points or nodes, together with a set of branches, each of which connects a pair of nodes (Myerson 1991)

is justified. In strategic form models one focal point is to measure the value of a network and determine how the value is distributed amongst the players. Random models define the characteristics of a network with several indicators that I thus illustrate here.

2.1 Overview of the literature and terminology

The first papers in the field of networks in economics published in 1950's studied coalitions instead of networks; e.g. Shapley (1953). To be precise, endogenous coalition formation was implicit already in the von Neumann-Morgenstern (1944) theory of stable sets. The models of network formation have evolved from the models of coalition formation which makes understanding the notion of coalitions important. The properties of coalitions and networks have some different features; coalition structure is among Aumann and Myerson (1988) the simplest model of "framework of negotiations". Coalition structure is defined as a partition of the player set into disjoint coalitions. Once the coalition structure has been determined, negotiations take place only within each of the coalitions that constitute the structure; each such coalition B divides among its members the total amount v(B) that it can obtain for itself.

Myerson (1977) thought that coalition structure is not rich enough to capture the subtleties of negotiation frameworks; he referred to networks as cooperation structures[3], whereas much of the literature that followed Myerson has used the term communication structures. Subsequently, Jackson (2003) called them communication games. Network games include cooperative games and communication games as special cases, but generally network games allow for costs and benefits to accrue differently to different sets of links, and allow for externalities and such across players and networks (Jackson 2003). Van den Nouweland (2004) referred to networks as communication networks which describe the bilateral channels through which individuals can communicate and coordinate their actions.

There are two alternative approaches to model the network formation. One derives from random graph theory and views an economic or social relationship as a random variable. The other regards people or firms or other actors involved as exercising discretion in forming their relationships, and uses game theoretic tools to model formation. (The New Palgrave Dictionary of Economics) In the basic random graph theory (e.g. Erdös and Rényi 1959) links form independently and the probability of link formation is distributed among Poisson distribution. In the study random graph models the emphasis is on examining the structure and properties of the graph. The strategic form network

[3] A cooperation structure is a graph whose vertices are identified with the players. A link between two players means that these players can carry on meaningful direct negotiations with each other. Coalition structure is a cooperation structure where two members i and j are linked iff they are in the same coalition (Dutta et al. 1998).

formation games describe the actual formation process. These models can be further divided into strategic models (e.g. Jackson and Wolinsky 1996) and game-theoretic models of network formation. There are both noncooperative network formation (e.g. Bala and Goyal 2000) and cooperative network formation models (e.g. Aumann and Drèze 1974) in the game theoretic approach. Often, the cooperative games are presented in reduced form as a noncooperative game due to simplicity. The difference between strategic models and game theoretic models underlies in their concepts of equilibrium.

2.2. Modeling networks

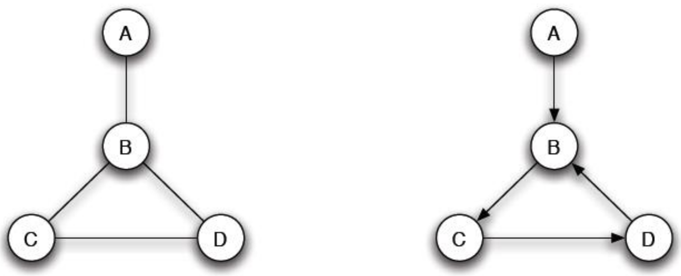
The set N = {1, ..., n } is the set of nodes that are involved in a network of relationships. Nodes are referred to as vertices, agents or players; nodes can be e.g. individual people, firms, countries or other organizations. (Jackson 2008a, Jackson and Wolinsky 1996) The notation ij Є g indicates that i and j are linked in network g. G = {g C g^N} is the set of all possible networks on N, g^N is thus denoted as the complete network (Marco Marini 2007). A graph (N,g) consists of a set of nodes N = {1, ..., n } and a real-valued nxn matrix g, where g_{ij} represents the relation between i and j (Jackson 2008a).

For curiosity, a game in coalitional form is denoted as follows. Let N = {1, ..., n } be the set of players and (N,v) the game in characteristic function form. Coalition structure B = {B_1, B_2,… B_m} is defined to be a finite partition either of the player set N or of the universe of v. Thus, game in coalitional form is denoted by (N, v, B).

An important goal in the analysis of strategic networks is to evaluate the economic welfare of an emerged network to the society, thus one has to be able to calculate a value to a network (Marini 2007). With a value function, one can assign each network a worth. Formally, a value function for a network is stated as $v: G \rightarrow R. V$ is the set of all possible value functions. Also, the characteristic function $V: 2^N \rightarrow \mathbb{R}$ describes how much collective payoff a set of players can gain by forming a network (Myerson 1991). However, value function is a richer object than a characteristic function of a cooperative game because the calculation may involve both costs and benefits as it allows the value that accrues to depend on the network structure (Jackson 2003). While the value function measures the worth of the whole network, the distribution of value is indicated by an allocation rule. An allocation rule is a function $Y: G \times V \rightarrow R^N$. Here, $Y_i(g, v)$ is the payoff to player i from graph g under the value function v. (Jackson and Wolinsky 1996) Allocation rule is important not only due to fairness considerations – the fact that value would be divided fairly between the players in the network – but also because allocation rule determines players' incentives to form various networks

(Jackson 2003). The Shapley value is probably the most prevalent allocation rule in the literature[4]. These concepts provide a useful basis for making predictions about the outcomes of multilateral bargaining on networks, as well as more generally for analyzing the power or influence of various players in a network (Jackson 2008). I will not focus on stability and efficiency issues of networks in this work, and therefore will leave the value functions and allocation rules aside.

The most common network form to appear in the strategic network formation models is the undirected graph, where two nodes are either connected or they are not. An undirected graph is such that one node cannot be related to the second without the second being related to the first (Jackson 2008a). A nondirected network (N,g) thus describes a system of reciprocal relationships between individuals (Marini 2007). For instance, if a network is a social network of people and links represent friendships or acquaintances, then it would tend to be non-directed. This network structure usually dominates in many social and economic relationships, such as partnerships, friendships and alliances. In Picture 1, the left-hand side network is an example of an undirected graph and the right-hand side of directed one. The arrows point out the direction of the link, from which the utility is retrieved. Here, A is connected to B but B is not connected to A. Some structures are better modeled as directed networks, where one node can be connected to second without the second being connected to the first. Directed graphs are found in networks that keep track of which authors cite which authors, which web pages have links to which others and in social networking sites where one can follow a user without his/her consent such as Twitter. The networks is directed if it is possible that $g_{ij} \neq g_{ji}$, and a network is undirected $g_{ij} = g_{ji}$ for all nodes i and j. (Jackson 2008)



Picture 1. Undirected and directed network. (Easley and Kleinberg 2009)

A length of a path is defined as the number of steps it contains from beginning to end, i.e. the number of edges in the sequence that comprise it. The paths tend to be surprisingly short and they connect a large fraction of the world's population when a global network of friends is considered.

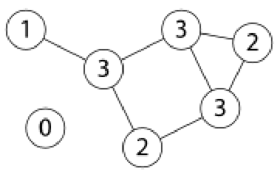
---

4 For more discussion about allocation rules and value functions refer to e.g. Shapley (1953), Slikker and van den Nouweland (2001), Myerson (1977), Myerson (1991), Aumann and Drèze (1974), Harsanyi (1963), Aumann and Myerson (1988).

These features of short path length and the existence of a giant component apply to a large fraction of networks; the phenomenon is known as the small-world phenomenon (Easley and Kleinberg 2009).

Random graph networks are described by a few straight-forward concepts. *A path* is simply a sequence of nodes with the property that each consecutive pair in the sequence is connected by an edge. A path not repeating the nodes is called a simple path. An important non-simple path is a *cycle* with at least three edges in which the first and the last nodes are the same. A graph is *connected* if for every pair of nodes, there is a path between them. If a graph is not connected, then it breaks apart naturally into a set of connected “pieces," groups of nodes. A connected component of a graph is a subset of the nodes such that: (i) every pair of nodes in the subset has a path to every other; and (ii) the subset is not part of some larger set with the property that every node can reach every other. Moreover, large, complex networks often have what is called a giant component, an informal term for a connected component that contains a significant fraction of all the nodes. (Easley and Kleinberg 2009)

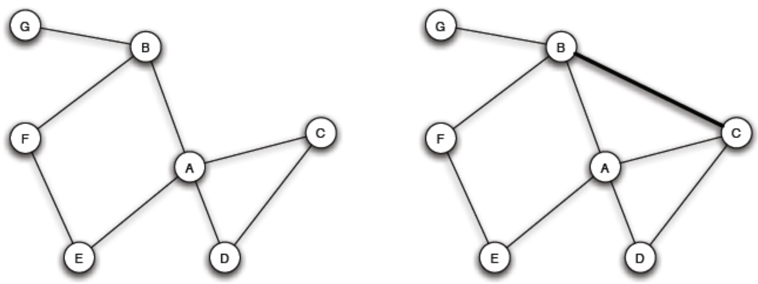
Most of the papers addressed to study random graph models actively evaluate the fit of emerging networks with real social networks. Here are a few measurements employed in the comparison. The degree of a vertex or a link in a graph is the number of edges that connect to it. The maximum and the minimum degree of a graph are simply the maximum and the minimum degree of its vertices (Jackson 2008b). In picture 2, the maximum degree is 3 and minimum degree is 0. In a regular graph, all degrees are the same where there is only a single degree of the graph.



Picture 2. The degree of a graph. (Wikipedia)

Regular graphs are rarely observed; the extent to which the degrees of graphs differ is measured by degree distributions. There are two important facts about the degree distribution properties of networks. Firstly, many social networks exhibit fat tails in that there are more nodes with relatively high and low degrees than would tend to arise if links were formed independently and totally at random. Secondly, it is hard to find networks that would follow a strict power law. The term power law refers to the frequency of a given degree being proportional to the degree raised to a power. (Jackson 2008b)

Clustering occurs in a network if the probability of two vertices being connected by an edge is higher when the vertices in question have a common neighbor, i.e. there is another vertex that they are both connected (Newman 2002), is also characteristic of networks. Besides degree distribution properties, also clustering properties, i.e. how clustering is distributed across of a network gives useful insight in the analysis of networks. Clustering is measured by a clustering coefficient which is defined as the probability that two randomly selected edges of a player are in fact linked with each other. As the clustering and properties of degree distributions imply, and as I will later show, links may not form independently. There are several factors affecting link formation which also explain the clustering of the nodes; many of the links form among the principle which is known as the triadic closure: *If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future*. (Easley and Kleinberg 2009).



(a) *Before B-C edge forms.* (b) *After B-C edge forms.*

Picture 3 Triadic closure. The formation of the edge between B and C illustrates the effects of triadic closure, since they have a common neighbor A. (Easley and Kleinberg 2009)

There are numerous underlying factors affecting the mechanism of the formation of social networks that also give rise to the phenomenon of triadic closure. Easley and Kleinberg (2009) refer to factors influencing network formation as network's surrounding contexts: factors that exist outside the nodes and edges of a network, but which nonetheless affect how the network's structure evolves. One of the most basic notions influencing the structure of social networks is homophily which here means that individuals tend to have links to others who are similar to them.

Furthermore, the links formed do not need to be the same intensity. Usually, the intensity or the strength of a link is interpreted so that stronger links represent closer friendship and greater frequency of interaction. Granovetter (1974) categorize the links in strong and weak ties. For research purposes one should distinguish between different levels of strength in the links of a social network although measuring the link strength might be quite complicated in practice. The strength of a link might have a significant impact e.g. to the way in which information between edges flows thus it is an important aspect to take into consideration.

Now, I will move on to discussing the nonstrategic, random graph models of network formation.

3. Nonstrategic models of network formation

From the two groups of network formation, nonstrategic models view networks as arising stochastically and therefore use random graph theory whereas strategic models see links in a network as social or economic relationships chosen by the players involved, the approach thus being game theoretic (Jackson 2008b). In this part, I present the basic random graph models. There are also many papers discussing the extensions of the basic models, that strive for replicating the properties of real networks. The need to move towards subtler models emerges when the social connections' dependence on existing relations is taken into account in the modeling process. I examine static extensions with preferential attachment properties, for instance (Barabási and Albert 1999). Dynamic modifications that I study develop the idea of preferential attachment further by assuming not self-reliant link formation (e.g. Golub and Jackson 2008).

Random graph models are useful in modeling phenomena like the spread of information or decisions that are heavily influenced by peers and in modeling the diffusion of diseases, for example. (Jackson 2008a). Moreover, Granovetter (1974) speaks about the importance of weak ties in e.g. labor markets. Among him weak ties or links are more relevant in finding out about new job opportunities than strong ties. Although the strength of the link and the degree of randomness of the formation process is not explicitly discussed in the literature, I believe that randomly generated links represent a weaker form of links than the ties which form through the strategic decision making. Also, it might be that job information networks – the networks through which job information flows – for the most part form among a random process.

3.1. Random-Graph Models of Network Formation

A random graph consists of some number N of nodes, connections or edges are placed between them so that each pairs of nodes i,j has a connecting edge with independent probability p (Newman et al.2002). After presenting the Erdös-Rényi model, I deliberate how the properties of a random graph and real networks differ. Thereafter I speak about the extensions of Erdös-Rényi model.

3.1.1. Basic Random-Graph Models

The Poisson random-graph models are one of the most extensively studied among the static models, the model of Erdös and Rényi (1959) being one of the first and the most "popular" Poisson models. Static refers to a typed model in which all nodes are established at the same time and the links are drawn between them according to some probabilistic rule (Jackson 2008a).

The random graph model of Erdös and Rényi (1959) consists of n nodes or vertices that are joined by links or edges, pairs of vertices are chosen uniformly at random. There are two closely related variants of Erdös-Rényi-model (1959) of which more commonly studied model is G(n,p) where each link is formed with probability p, $p \in (0,1)$, and the average number of links or edges in the graph is $pn(n-1)/2$ (Erdös and Rényi 1959, Newman 2002). The average degree z of a vertex or a node is

$$z = \frac{n(n-1)p}{n} = (n-1)p \simeq np$$

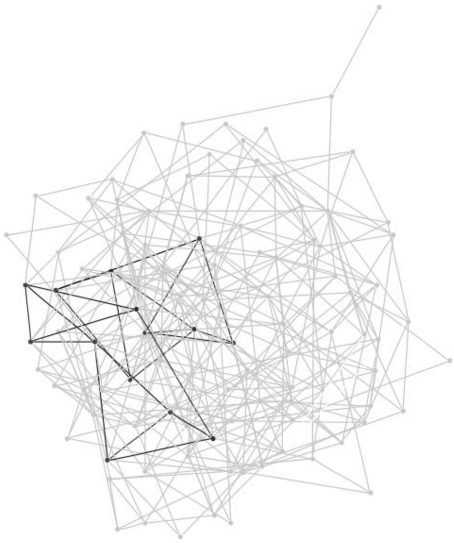
for large enough n. The distribution of the degree of a vertex is binomial:

$$p_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

where k is the degree of the node. When the equation becomes

$$p_k = \frac{z^k e^{-z}}{k!}$$

which is the Poisson distribution with large enough n.



Picture 4. **Erdös-Renyi Random Graph G, edges=316, nodes=100. (www.math.ucsd.edu/~fan/algo/sigma.html)**

With increasing degree distribution the Erdös-Rényi model shows phase transition which is when the giant component forms. A component is a subset of nodes in the graph each of which is reachable from the others by some path through the network. There is a critical value of degree distribution z above which the one largest component in the graph contains a finite fraction S of the total number of vertices. The phase transition occurs at z=1. The largest component is called the giant component. Goyal, van der Leij and Moraga-Gonzáles (2006) say that the network has a giant component if the largest component constitutes a relatively large part of the vertices and all other components are small, typically of order $\ln(n)$. (Newman 2002)

3.1.2. Properties of random-graph networks and real social networks compared

There are some features that random-graph models lack compared with real-life social networks. This is mainly because links do not necessarily form independently in reality. I will discuss dependencies in random network formation in the next part and in part 3.2.5. Firstly, real-world networks show strong clustering or network transitivity whereas Erdös and Rényi's (1959) do not. Secondly, they exhibit "small-world" effect (Newman et al.2002). Thirdly, social networks tend to have skewed degree distributions vis-à-vis the random graph models; the degree distribution of real networks is thus highly different from the Poisson distribution (Newman et al. 2002, Newman 2002).

The "small-world" effect was introduced by Stanley Milgram (1967), in the experiment of his including letters. He concluded that many pairs of seemingly distant people are connected by a very short chain of intermediates; the typical length of chain is only about six[5].
Clustering can be measured e.g. by clustering coefficient C (Watts and Strogatz 1998). A clustering coefficient for random graphs is simply C=p (Newman 2002).

Below in table 1 I show a few examples how measured and random graph clustering coefficients differ:

Network	n	z	C measured	C random graph
Neural network	282	14.0	0.28	0.049
Power grid	4941	2.7	0.080	0.00054
Company directors	7673	14.4	0.59	0.0019

Table 1. **Measured and random graph clustering coefficients.**

where z is the average degree of a vertex in a network. The first two observations come from a paper by Watts and Strogatz (1998) and the bottom number from Newman et al. (2001). It seems that the less randomly the network is formed the less correlated are the clustering coefficients. I arranged the networks from the most random to least random, assuming that here neural networks would be the most randomly formed. In the neural network, the gap between measured and random graph

[5] This is also known as the Six Degrees of Separation (Newman et al. 2002)

coefficient C is the smallest, the difference being “only” 0.231. Nonetheless, the difference between the measured and random graph clustering coefficients is very significant, as well as is the case with degree distributions of the real networks and random graphs (Newman 2002).

3.2. Extensions of the basic random-graph models

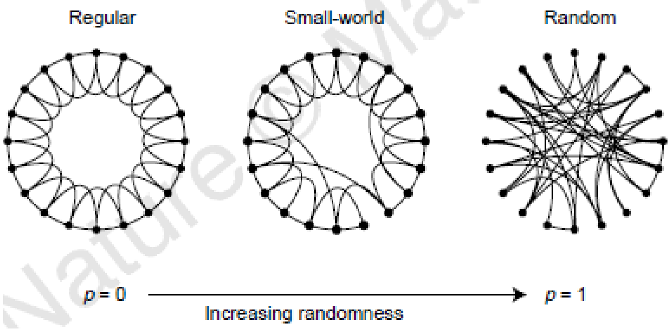
The basic random models capture the main idea of random network formation well but there are many shortcomings in the basic models and phenomena that random models fail to explain. The key motivation for the search of a well-fitting model is to provide a model that can be statistically estimated and still allows for specific dependencies between the probabilities whereby different links form (Jackson 2008b). Moving from static models to stochastic ones allows capturing both the regularities in the processes giving rise to network ties while at the same time recognizing that there is variability that one is unlikely to be able to model in detail (Robins et al. 2007). The network is usually conceptualized as a self-organizing system of relational ties. Substantively, the claim is that there are local social processes that generate dyadic relations, and that these social processes may depend on the surrounding social environment (i.e. on existing relations). For example, it is assumed that actors with similar attributes are more likely to form friendship ties (homophily), or that if two unconnected actors were connected to a third actor, at some point they are likely to form a friendship tie between them (transitivity) (Robins et al. 2007).

There are a few modeling alternatives for the Erdös-Rényi (1959); e.g. Newman et al. (2002) fix the degree distribution of a random graph model in order to mimic the clustering and the degree properties of real-world networks, others include Barabási-Albert (1999) dynamic scale-free and Watts-Strogatz (1998) small world model. The basic random graph model of Erdös and Rényi (1959) is static; while all the variations of the basic model that I refer to exhibit dynamic features. Watts and Strogatz (1998) model of clustering and small-world phenomenon illuminates the dynamics as an explicit function of structure rather than for a few particular topologies. Also, Barabasi and Albert (1999) assume a nonstatic network with a growing number of vertices and preferential attachment. To account for the above mentioned homophily, Golub and Jackson (2009) have created a probabilistic multi-type random network. Transitivity, or more generally the assumption that nodes do not form independently is incorporated in the exponential random graph models, known as p*models, e.g. Wasserman and Pattison (1996).

3.2.1. A small-world model

Watts and Strogatz (1998) generate a small-world model, a variation of a random network, where they show that only a small number of random links in a network are needed to generate a small diameter. They define small-world networks as having short paths and high clustering. Among Goyal et al. (2006) network G exhibits small-world properties if: (1) the number of nodes is very large compared with the average number of links, n>n(G) (2) the network is integrated; a giant component exists and covers a large share of the population. (3) The average distance between nodes in the giant component is small. (4) Clustering is high. Property (3) refers to short paths.

Watts and Strogatz (1998) explore simple models where regular, nonrandom, networks are rewired to introduce increasing amounts on randomness, disorder. The rewiring procedure starts from a ring lattice with n vertices and k edges per vertex, then each edge is rewired at random with probability p. Watts and Strogatz (1998) quantify the structural properties of the graphs by their characteristic path length L(p) and clustering coefficient C(p). L(p) measures the typical separation between two vertices in the graph whereas C(p) measures the cliquishness of a typical neighborhood. A nonrandom regular lattice at p=0 (at p=0 network is regular and at p=1 it is completely random) is highly clustered large world where L grows linearly with n, whereas the random network at p=1 is poorly clustered, small world where L grows only logarithmically with n.



Picture 5. Random rewiring procedure for interpolating between a regular ring lattice and a random network, without altering the number of vertices or edges in the graph. (Watts and Strogatz 1998)

The small-world networks result from the immediate drop in L(p) caused by the introduction (as a result from the rewiring process) of a few long-range edges. These “short cuts” connect vertices that would be much farther apart in a random network. The important implication in Watts and Strogatz’s (1998) paper is the fact that an edge removed from a clustered neighborhood to make a short cut doesn’t change C(p) even though L(p) drops rapidly. Thus the transition to a small world is almost inevitable. Small-world phenomenon might be common in sparse networks with many vertices (Goyal et al. 2006) property (1) as only few short cuts are needed to make the network a small world. In a small world network information or diseases, for example, are predicted to spread much more easily and quickly. Random small-world networks allow for the quick spread of different flows, which could be thought as random phenomena, but the random small world network structure is much less suitable for enforcing cooperation between players. Watts and Strogatz (1998) find that for the multi-player "Prisoner's dilemma" played on a graph, as the fraction of short cuts increases, cooperation is less likely to emerge in a population of players using a generalized "tit-for-tat" strategy[6]. The likelihood of cooperative strategies evolving out of an initial cooperative/non-cooperative mix also decreases with increasing p. This result is quite reasonable; as the randomness in a network increases the possibility for cooperative strategies for emerge decreases. Easley and Kleinberg (2009) propose that since the randomly connected nodes are completely unrelated assuming a heterogeneous network, they are hard for agents to use reliably. In fact, no trust between agents is likely to emerge. It seems that the structure of random networks enables poorly for cooperation, the social or economic networks allowing for cooperation are most probably likely to form strategically.

3.2.2. Fixed degree distributions

Another variation of a Poisson distributed random graph model which tries to account for the nonrandom social phenomena is a model by Newman et al. (2002) where the degree distribution is fixed. They are thus given the probabilities p_k that a randomly chosen vertex in the network has a degree k. By fixing the degree distributions, one can avoid the problems with different degree properties of random and real networks. They take a number of vertices and assign to each a number k of "stubs" or ends of edges, where k is a random number drawn independently from the distribution p_k for each vertex. The stubs are chosen randomly in pairs and joined up to form edges between the vertices. This will produce a graph with exactly the desired degree distribution. In calculating the properties of the emerged network Newman et al. (2002) apply a generating function $G_0(x)$ instead of using the degree distribution p_k which is defined as

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k.$$

The model networks' most striking property is that they exist in two different regimes. Depending on p_k the networks either can consist of small clusters of vertices connected together by edges – components - or they may contain a giant component. If there is no giant component in a network, then all components are small and communication can only take place within small groups of people of typical size (s). If a giant component does exist, then a large fraction of the vertices in the network all can communicate with one another (and the number S is this fraction). (Newman et al. 2002) The degree distribution properties of networks thus determine how the information moves within a

[6] A tit-for-tat strategy plays the iterated prisoners' dilemma game by cooperating on the first move, and then making the same choice as the other player did on the previous move (Routledge Encyclopedia of International Political Economy 2001, edited by R. J. Barry). It was first introduced by Anatol Rapoport and Robert Axelrod aroun 1980.

network. Among Newman et al. (2002) almost all networks in society and nature seem to have a giant component; networks with no obvious giant component are rare.

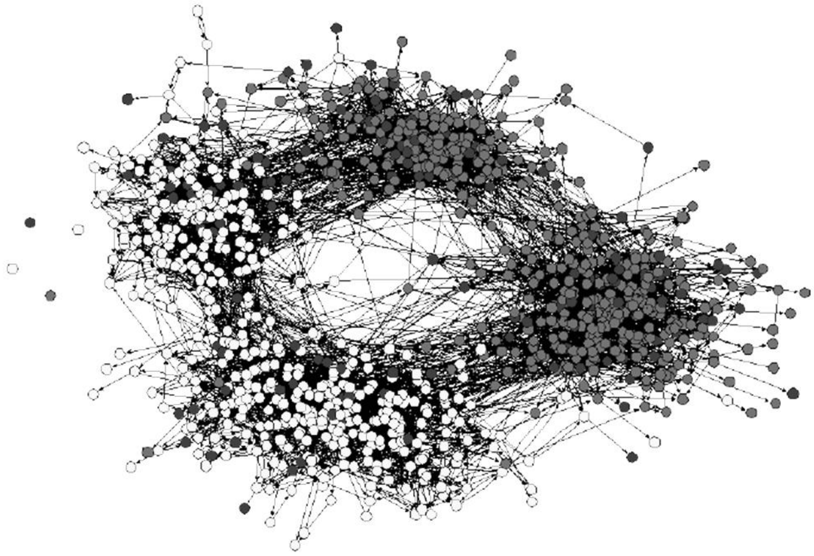
3.2.3. Preferential attachment

Most of the existing random graph models fail to incorporate growth and preferential attachment. To eliminate the usual shortcomings, Barabási and Albert (1999) have created a model that generates scale-free networks using the assumption of preferential attachment and growth. In many large networks, the vertex connectivities follow a scale-free power-law distribution. The scale-free feature is due to the fact that networks expand continuously by the addition of new vertices, and new vertices attach preferentially to already well connected vertices. The initial difference in the connectivity between two vertices will thus increase further as the network grows. The older (smaller t_i) vertices increase their connectivity at the expense of the younger (larger t_i) ones, eventually leading to a setting where some vertices are highly connected. This kind of "rich-gets-richer" phenomenon can easily be found in real networks. Barabási and Albert (1999) study www sites and citation patterns and they show that the probability P(k) that a vertex in the network interacts with k other vertices decays as a power-law, following $P(k)\sim k^{-\gamma}$. This indicates that large networks self-organize into a scale-free state, a feature that is not present in any existing random network models. Also, in most real world networks the number of vertices increases throughout the lifetime of the network. In Erdös and Rényi (1959) and Watts and Strogatz (1998) model the probability of finding a highly connected vertex, a vertex with large k, decreases exponentially with k; vertices with large connectivity are nearly nonexistent. In contrast, the power-law tail characterizing P(k) in Barabasi and Albert (1999) indicates that highly connected vertices have a large chance of occurring.

3.2.4. Homophilous networks

One major shortcoming in the network literature is that most of the random and strategic form models fail to account for the pervasive fact that links in a social network tend to connect people who are similar to one another. One exception is the paper of Golub and Jackson (2008) who study how information transmission in various setting is affected by homophily[7] which is observed almost universally in social networks. Homophily is defined as the tendency to associate with others with similar characteristics such as race, age, gender, and profession, religion and various behaviors (Golub and Jackson 2008) of which the latter three are defined as choice homophilous characteristics.

[7] See Currarini, Jackson and Pin (2007) for a friendship formation model based on a matching process that also introduces homophily.



Picture 6. **Homophily can produce a division of a social network into densely-connected, homogeneous parts that are weakly connected to each other. Here, the different shades represent different races (Easley and Kleinberg 2009).**

A high degree of homophily in a network can have various consequences; especially affect the transmission of information. The influence mechanism of homophily in a network is for instance such that the information individuals learn from same-type individuals is more likely to be correlated with their own information, and that of different types is more likely to be uncorrelated. While the information flowing in a heterogeneous network might be more valuable, a same-type individual is easier to communicate with. This applies also to professional networks; individuals alike are easy to communicate with, but offer less creative synergy (Currarini et al. 2007). Consequently, homophily might in effect result in lower productivity and welfare.

Golub and Jackson use a probabilistic multi-type random network model which generalizes the Erdös-Rényi random network model. Here, agents are divided into different types and the links are formed independently between various agents, with the probability of a link forming between two agents depending on the types of agents involved. They consider three different processes where the homophily in the network structure affecting diffusion turns out to depend on the type of learning or diffusion process. The two latter are affected by homophily whereas the first process is not. The first is the shortest path process where information is navigated to its destination via the shortest path, examples include many peer-to-peer systems like Internet and information spreading phenomena. The second is based on linear updating or learning where individuals update their beliefs or actions by repeatedly taking weighted averages of their neighbors' beliefs or actions. The third, and the most interesting from my perspective, is a random walk process on a network, where some particle hops around the network, having equal probability of moving along any link out of its current node. The third process could fit well in modeling the diffusion of job information in reality where the particle could be thought e.g. as job information.

Here, the multi-type random network incorporating homophily is presented as follows. A network is represented via its adjacency matrix: a symmetric n-by-nmatrix **A** with entries in {0,1}, given the set of nodes $N = \{1, \dots, n\}$. The interpretation is that $A_{ij} = A_{ji} = 1$ indicates that nodes i and j are linked, and the attention is restricted to undirected networks. Let $d_i(\boldsymbol{A}) = \sum_{j=1}^{n} A_{ij}$ denote the degree of node i. Let $d_{min}(\boldsymbol{A})$ and $d_{max}(\boldsymbol{A})$ be the minimum and maximum degrees, respectively and $\bar{d}(\boldsymbol{A})$ denote average degree, and let $D(\boldsymbol{A}) = \sum_i d_i(\boldsymbol{A})$ be the total degree in the society. The actual multi-type random network consists of a vector $\boldsymbol{n} = (n_1, \dots, n_m)$ which captures how many nodes of each type there are (and how many types, m, there are), and a symmetric m-by-m matrix **P**, whose entries in [0,1] describe the probabilities of links between various types. Let N_k be the set of nodes of type k, and label nodes so that $\{1, \dots, n_1\}$ are the nodes of the first type, $\{1 + n_1, \dots, n_1 + n_2\}$ are the nodes of the second type, and $N_k = \{1 + \sum_{i<k} n_i, \dots, \sum_{i \leq k} n_i\}$ are the nodes of the k-th type. The resulting random network is captured via its adjacency matrix which is denoted **A**(**P**, **n**) and is a random variable. **A**(**P**, **n**) is built by letting the entries A_{ij} with be independent Bernoulli random variables with parameter P_{kl} if $i \in N_k$ and $j \in N_l$. That is, the entry P_{kl} captures the probability that an agent of type k links to an agent of type l.

A special case of the model is an island model which also illustrates the random walk and linear updating model; this model is able to demonstrate the influence of homophily on the emerged networks more clearly. Thus, the nodes of the same type connect to each other with one probability and nodes of different types connect to each other with another probability.

Golub and Jackson define two measures of homophily; the (unnormalized) homophily is defined as $H = \frac{p_s}{p}$ which captures how much more probable a link to a node of one's own type is compared with other types. This varies between 0 and m, the number of islands. If a node only links to same-type nodes, then the average linking probability p becomes ${p_s}/{m}$ and so $H = m$. If a node only links to the nodes of other types, then $p_s = 0$ and naturally $H = 0$. The normalized homophily, where the measure is divided by the number of islands m, is defined as $h = \frac{p_s}{mp}$. Thus, h is the fraction of a node's links that are expected to be the agents of the same type.

Golub and Jackson show that homophily affects the random graph and averaging process by possibly substantially slowing the process down whereas the shortest path process is unaffected. The slowdown is due to the fact that even though the average path length is unchanged, there are relatively fewer paths between the agents of different types as homophily increases. This result indicates that homophily could have a significant effect also in the spread of job information within the network. If a network consists of homophilous components, the diffusion of information between the components would be very slow or non-existent which means that not everyone even inside the network is eligible to certain information. This could very well explain the “rich-gets-richer” phenomenon. All in all, random models incorporating homophily find that there are both chance and choice involved in network formation. This approach is clearly a step away from complete randomness in network formation.

3.2.5. Exponential random graph models

Development of exponential random graph models is another approach to correct the assumption of independent link formation of the basic Bernoulli random graph distribution proposed by Erdös and Renyi (1959). The exponential models are known as p* models. The independency of edges is clearly an implausible assumption in almost all human social networks, but there are many factors affecting the pattern of link forming. Transitivity is one of the outstanding features that differentiate observed data from a pattern of random ties. It is expressed by a triadic closure; when there is a tie from i to j, and also from j to h, then there is a tie from i to h: “*friends of my friends are my friends*” (Snijders et al. 2006). There are several social processes which may give rise to transitivity. First, social ties may self-organize to produce triangular structures just as happens in the process shown above in Chapter 2. Alternatively, certain actors may be very popular, and hence attract ties as is the case in the preferential attachment model of Barabási and Albert (1999). This process may result in core periphery network structure with popular actors in the core. In a separate third possibility, however, ties may arise because actors select partners based on attribute homophily, in which case triangles of similar actors may be a byproduct of a homophilous dyadic selection process. For a full understanding about the processes that give rise to and sustain the network, it is crucial to model transitivity adequately. Exponential random graph models offer the most promising framework within which such models including transitivity can be developed because these models allow for specific dependencies between the probabilities whereby different links form. (Snijders et al. 2006)

To demonstrate conditional dependencies, Markov graph models have been introduced to real social networks. A Markov graph[8] is a generalization of Poisson random graphs that has been useful in the

[8] Markov graphs derive from Markov chain. A Markov chain is a sequence of random variables X_1, X_2, X_3…(all state transitions are probabilistic) with the Markov property that, given the present, the future is conditionally independent of the past. I.e. the description of the present state fully captures all the information that could

statistical analysis of observed networks. A Markov graph is also one form of exponential random graph but these models turn out to be not flexible enough to represent the degree of transitivity observed in social networks. The inflexibility stems from the fact that the Markov dependence assumption of Frank and Strauss (1986) is too strong and less strict conditional independence assumptions must be made (Snijders et al. 2006). In addition, the Markov dependence seems unrealistic for large networks, where individual actors may not even be aware of each other and have no means to come into contact, yet their possible tie still is taken to influence other possible ties. Exponential random graph models that go beyond Markov random graphs include Wasserman and Pattison (1996), Pattison and Robins (2002) and Snijders (2002)[9].

Conditional dependencies can be established so that the probability of a link *ik* depends on whether *ij* and *jk* are present. Such dependencies tend to interact with one another in ways that could make it impossible to specify the probability of different graphs in a tractable manner (Wasserman and Pattison 1996). In positing an exponential random graph model five steps are implicitly followed. In step 1, each network tie is regarded as a random variable. In the model there is random variable *Yij* where *Yij* = 1 if there is a network tie from actor *i* to actor *j*, and where *Yij* = 0 if there is no tie. Notation *yij* is specified as the observed value of the variable *Yij* and **Y** is the matrix of all variables with **y** the matrix of observed ties, the observed network. In step 2, a dependence hypothesis is proposed, defining contingencies among the network variables. This hypothesis embodies the local social processes that are assumed to generate the network ties. In step 3, the dependence hypothesis implies a particular form to the model; well-specified dependence assumptions imply a particular class of models. In step 4, parameters are simplified through homogeneity or other constraints in order to reduce the number of parameters to be able to define the model more clearly. Finally, in step 5 model parameters are estimated and interpreted.
The general form of the exponential random graph model is

$$P(\boldsymbol{Y} = \boldsymbol{y}) = \frac{1}{\boldsymbol{k}} \exp\{\sum_A \eta_A g_A(\boldsymbol{y})\}$$

---

influence the future evolution of the process. The transition probabilities of moving from a one state to another are dependent only upon the starting state of any transition, rather than upon how that state was reached (Diaconis 2008) . Formally, $\Pr(X_{n+1} = x | X_n = x_n, \ldots, X_1 = x_1) = \Pr(X_{n+1} = | X_n = x_n)$. (http://mathworld.wolfram.com/MarkovChain.html/)

[9] See Robins et al.(2007) and Snijders et al. (2006) for more discussion.

where ηA is the parameter corresponding to the configuration $A, g_A(\boldsymbol{y}) = \prod_{y_{ij\epsilon A}} y_{ij}$ is the network statistic corresponding to configuration A and κ is a normalizing quantity which ensures that the equation is a proper probability distribution.

3.3. Conclusion

Poisson distributed static random graph models are a simple way to model theoretical random networks which do not demonstrate dynamic features. The model of Erdös and Rényi (1959) is perhaps most often referred to in the literature. In this model, nodes are connected by links independently with a probability p. The basic random graph models fail to demonstrate some features typical for real life social networks. Real life networks show strong clustering, a “small-world” effect and they tend to have skewed degree distributions compared with the networks generated by random models. Therefore, there have been many attempts to improve the fit of the static random graph model for real networks in the literature.

To sum up; a typical variation of the random graph model is a model with fixed degree distributions. By fixing the degree distribution this kind of model tries to mimic the degree distribution properties of the real networks. Newman et al. (2002) conclude that depending on the distribution their model networks can either consist of small clusters or they may contain a giant component. Even though in Newman et al. (2002) the degree distribution properties are the same as real networks they still are not able to model networks which would always show “small-world” properties. It must be that imitating degree properties does not substantially improve the fit of the model networks compared with real networks. Better fit is obtained through introducing dynamics, i.e. stochastic models.

In Watts and Strogatz (1998) one of the key findings is that the transition to small-world networks is almost unavoidable. Small world networks have short paths and high clustering. With their dynamic rewiring process they have been able to better model the existing networks. The small world model of Watts and Strogatz (1998) illustrates dynamics as an explicit function of structure. The dynamics is present also in Barabási and Albert (1999) model, here; it is based on the assumption of growing networks (growing number of vertices) and preferential attachment. Thus, this model should avoid the shortcomings of the basic random graph models. The networks generated by the models feature scale-free properties, i.e. networks expand by the addition of new vertices and new vertices attach to already well connected vertices. Contrary to Newman et al. (2002) and Watts and Strogatz (1998) models, highly connected sites are very likely to be found.

Nonetheless, more often than not the links between nodes do not form independently which the basic random graph model does not account for, which is why the models do not manage to reproduce the characteristics of real networks. Golub and Jackson (2009) present the assumption of preferential attachment more explicitly by studying the effects of homophily in networks. Also, transitivity is another key feature present in real social networks which impacts greatly the structure of social networks.

The Markov graph model allows for dependencies in real life networks but it still seems that the more nonrandom and more dynamic network the less precisely the static random graph models are able to model these networks. Markov graph dependence is a step to the right direction in formulating network formation but it imposes too strict and unrealistic dependencies. Robins et al. (2007) and Snijders et al. (2006) speak about exponential random graph models incorporating transitivity that move beyond Markov dependencies. Compared with other models presented here; these three papers assume conditional probabilities instead of total randomness, since both choice and chance that matters in network formation. Properly formulated assumptions enable including the choice also in the random models. Indeed, the merit of the papers is that they acknowledge that links do not form independently and develop the dependence assumptions beyond Markov graph dependencies.

However, the fact that the formation of nonrandom networks is modeled with random graph models leaves the study of link formation somewhat one-sided. Random graph models and the variations do not model the network formation from the perspective of the nodes, i.e. players or agents, these models have more a planner's view. That is, what Newman et al. (2002) call nonrandom social phenomena, which they try to take into consideration in their work, could be thought as individual incentives, willingness to form a link. To truly include the "nonrandom social phenomena" one has to change perspective, which leads to a strategic approach. By changing the perspective, one can actually account for the individual incentives in network formation while random models only recognize the fact that some links are more likely to form than others. However, I do not suggest that strategic form models could water tightly describe how networks form in real life, but they still bring a new standpoint into the discussion. Also, the focus of strategic form models is not so much on reflecting how well the models fit real social networks but rather on examining the stability and the efficiency of the emerged networks. In addition, not all the job information is spread through weak links that are probably more likely to emerge through random models; therefore, strategic form models that I assume to produce stronger links should also be considered. Overall, it is worth noting that not all links form randomly and consequently, not all form strategically.

In the next part I first study the network formation models in strategic form.

4. Strategic and game theoretic modeling of network formation

Random graph models pursue to reproduce the properties of the real networks in the emerging model networks whereas the basis of strategic form models is the concepts of stability and efficiency. The structure of the real networks does not determine which links should form in the strategic form models, only the fact that a new link does not violate the conditions of stability and efficiency. Usually, the strategic approach models are categorized based on their notions of equilibrium; Jackson (2008a) calls the two approaches strategic and game theoretic models. The notion of pairwise stability is a feature of strategic models, whereas the basic Nash-equilibrium is present in the game theoretic models of network formation. The Nash equilibrium is more widely used even though pairwise stability would perhaps be more suitable in the context of networks.

The strategic form network formation models are well represented in the literature. I limit the discussion to two main models, the ones of Jackson and Wolinsky (1996) and Bala and Goyal (2000).

4.1. Comparison of the random-graph and strategic networks

The approach and the assumptions of random graph models are quite different from strategic form models. Contrary to strategic form models, random graph approach pay attention to the goodness of the fit so as to illustrate the existing social structure as well as possible. The papers about strategic form models concentrate on reflecting stability and efficiency of the network structure that the models produce, for instance, the whole model of Jackson and Wolinsky (1996) is based on a stability test. Essentially, the approaches being so different, it might be that random graph models actually have more realistic assumptions about social networks than what strategic models do. Even though random graph models describe the link formation on a network level well, strategic models better describe networks where economic agents want to maximize their utility and where the motivation to form a link with another agent consists in the link enabling for cooperation. The focus in strategic form models is on examining the link formation on a single agent level.

The difference between the two groups of network formation models could be expressed in terms of degree of activity of the agents in the formation process. As far as random graph models are considered, the links between nodes roughly just form without the consent of nodes or agents, they play a passive role in the process. Strategic models on the other hand demonstrate an active formation process, where forming or not forming certain links is a strategic decision for each utility maximizing[10], rational player forming them. That is, strategic form models enlighten the actual link formation process between individuals. As van den Nouweland (2004) defines, a network modeled by a strategic-form game[11] is a collection of bilateral links between players who must establish and maintain their own links. From the economic point of view, the focal point in strategic models is the study of link formation between individuals. Most networks might however form randomly, thus the structure of the economy might be fundamentally defined by the random formation processes. Therefore, it is justified to include both models in this work.

Nevertheless, when it comes to strategic form models, typically little is said about the factors affecting link formation, i.e. what are individuals' motivations to form or not to form a link, what underlying triggers such as transitivity in random graph models contribute to the formation of a link. Here, it is sufficient to understand the principle how links between two players form in theory and to know that there are factors (such as presented in Chapter 2) affecting link formation. In other words, while random graph models take account some of the underlying factors that affect network formation such as homophily, i.e. some of the recent models acknowledge that links do not form independently, strategic form models tend to leave these factors aside or take them as given and speak about the principles of how links between players actually form. The emerged networks are then subject to the analysis of stability and efficiency from the economic point of view.

4.1.1. Overview of the strategic form models

The literature of network formation in strategic form is quite extensive and exploring here even the majority of the models would be too ambitious and would not serve the needs of this paper. The strategic form models can be divided into cooperative and noncooperative approach models; the Jackson and Wolinsky (1996) being cooperative and Bala and Goyal (2000) a noncooperative model, I study both the models in this chapter. A cooperative game is a game where groups of players,"coalitions", may enforce cooperative behaviour. Nash (1951) defines cooperative game as a situation involving a set of players, pure strategies, and payoffs "as usual"; but with the assumption that the players can and will collaborate. An example is a coordination game when players choose the strategies by a consensus decision-making process. Myerson (1991) defines game to be

---

[10] Maximizing utility or wealth does not imply that the incentives are rational in terms of money value, for example, especially when social network formation is considered. Agents do not necessarily maximize utility by simply calculating the payoffs and costs of forming a link but utility is much more implicitly determined. (Jackson 2008)

[11] To define a game in strategic form, one needs to specify the set of players in the game, the set of options available to each player, and the way that players' payoffs depend on the options that they choose (Myerson 1991).

cooperative when players can negotiate effectively. In noncooperative games it is assumed that each player acts independently, without collaboration or communication with any of the others (Nash 1951).

To name a few earlier studies within this field, von Neumann and Morgenstern (1944) were the first to construct a cooperative theory of *n*-person games in coalitional form. Aumann and Drèze (1974) continue the coalitional approach and develop a solution concept generalizing several solution concepts developed earlier. Based on a cooperation structure introduced by Myerson (1977), Aumann and Myerson (1988) were the first to model network formation as a game in extensive form. Yet this model has not been widely used because it is not easily applicable beyond some simple examples.

4.2. Strategic network formation model

In recent years, different network formation procedures and network stability concepts have been proposed. There are principally three types of definitions of stability concepts, i.e. equilibrium: those definitions based on a pairwise stability notion, those based on the Nash equilibria of a link formation game, and those based on equilibria of a link formation game where transfers are possible. Here, I will be concentrating on the two first concepts of equilibrium. The whole picture of the different solution concepts is hard to decipher. For example, many models consider several notions of Nash equilibria in studying the possible equilibrium networks. Networks which are defined stable according to one solution concept are not stable among another concept; the connections between various solution concepts are rarely explicitly stated or even studied. (Bloch and Jackson 2006) I will analyze the issues of equilibrium, stability and efficiency in more detail later.

In their paper, Jackson and Wolinsky propose a stability test for social networks – pairwise stability – which is a notion that applies directly to the network and player's payoffs from networks (Bloch and Jackson 2006); the whole model of network formation builds on this stability test. In fact, the pairwise stability is introduced by Jackson and Wolinsky (1996). The main idea in Jackson and Wolinsky's (1996, 2006) models – the connections and the co-author model - is that the consent of both players is needed to form the link but a single player is able to sever the link. Jackson and Wolinsky (1996) have argued that the basic game theoretic model does not include communication and cooperation features but it may well be that an application of the basic model could be designed to take these features into account.

In Jackson and Wolinsky (1996) framework the value of a network can depend on exactly how agents are interconnected, not just who they are directly or indirectly connected to, thus the shortcoming of games with communication regarding to the network structure does not apply to the model of pairwise stability. The basic value functions and allocation rules are valid for the pairwise stability model. The value of a graph is represented by $v: \{g | g \subset g^N\} \rightarrow \mathbb{R}$. A graph $g \subset g^N$ is strongly efficient if $v(g) \geq v(g^{'})$ for all $g' \subset g^N$. The term strong efficiency indicates maximal total value. Strong efficiency and Paretian efficiency are equivalent if value is transferable across players. An allocation rule $Y: \{g | g \subset g^N\} \times V \rightarrow \mathbb{R}^N$ describes how the value associated with each network is distributed to the individual players. $Y_i(g, v)$ is the payoff to player i from graph g under the value function v.

The definition of a stable graph embodies the idea that players have the discretion to form or sever links. The formation of a link requires the consent of both parties involved, but the severance can be done unilaterally. Formally, this can be stated as follows: The graph g is pairwise stable with respect to v and Y if

(i) for all $ij \in g, Y_i(g, v) \geq Y_i(g - ij, v)$ and $Y_j(g, v) \geq Y_j(g - ij, v)$ and

(ii) for all $ij \notin g$ if $Y_i(g, v) < Y_i(g + ij, v)$ then $Y_j(g, v) > Y_j(g + ij, v)$

Condition (ii) embodies the assumption that, if i strictly prefers to form the link ij and j is just indifferent about it, then the link will be formed. The pairwise stability is a relatively weak notion among those which account for link formation and as such it permits a relatively larger set of stable allocations than might a more restrictive definition, such as Nash equilibrium, or an explicit formation procedure. However, a weak definition may provide strong results since it narrows the set of graphs substantially.

Next I discuss two stylized versions of the general model of Jackson and Wolinsky, the connections and co-author model. There are various versions of the basic model that one could think of but these two models are meant to capture the basic issues arising in social and economic network.

4.2.1. The Connections Model

This example models social communication among individuals. Individuals directly communicate with those to whom they are linked. They also benefit from indirect communication from those to whom their adjacent nodes are linked. The value of communication obtained from other nodes depends on the distance to those nodes. Jackson and Wolinsky (1996) also assume the communication to be costly thus the individuals must weigh the benefits of a link against its costs.

The utility of each player i from graph g is

$$u_i(g) = w_{ii} + \sum_{j \neq i} \delta^{t_{ij}} \, w_{ij} - \sum_{j\,:ij\,\in g} c_{ij}$$

where $w_{ij} \geq 0$ is the intrinsic value of individual j to individual i and c_{ij} is the cost to i of maintaining the link ij, t_{ij} is the number of links in the shortest path between i and j (setting $t_{ij} = \infty$ if there is no path between i and j), and 0<δ<1 says that the value that i derives from being connected to j is proportional to the proximity of j to i. To sum, less distant connections are more valuable than more distant ones, but still, direct connections are costly. Here, the value of a graph is

$$v(g) = \sum_{i \in \mathcal{N}} u_i(g)$$

In this model, the exact link formation process is not costly but maintaining the link incurs costs – in the case of a friendship the costs could consist of e.g. phone calls, travel costs, opportunity costs of the time consumed to maintaining the link.

By setting $c_{ij} = c$ for all ij and $w_{ii} = 1$ for all $j \neq i$ and $w_{ij} = 0$ one gets a symmetric model. The unique strongly efficient network in the symmetric connections model is

(i) the complete graph g^N if $c < \delta - \delta^2$,

(ii) a star[12] encompassing everyone if $\delta - \delta^2 < c < \delta + (\frac{N-2}{2}))\delta^2 \ and$

(iii) no links if $\delta + (\frac{N-2}{2})\delta^2 < c$

Jackson and Wolinsky also examine the stability for the allocation rule $Y_i(g) = u_i(g)$ without side payments. The specification of no side payments might correspond best to a social network where by convention no payments are exchanged for friendship. I do not see the stability with side payments important to consider here because I believe that majority of job information spreads in networks that resemble friendship networks. In the symmetric connections model with $Y_i(g) = u_i(g)$:

(i) A pairwise stable network has at most one (non-empty) component.

(ii) For $c < \delta - \delta^2$, the unique pairwise stable network is the complete graph, g^N.

(iii) For $\delta - \delta^2 < c < \delta$, a star encompassing all players is pairwise stable, but not necessarily the unique pairwise graph.

(iv) For $\delta < c$, any pairwise stable network which is non-empty is such that each player has at least two links and thus is inefficient.

For proof, refer to Jackson and Wolinsky (1996) p.51.

[12] The term star describes a component in which all players are linked to one central player and there are no other links: $g \subset g^N$ is a star if $g \neq \emptyset$ and there exists $i \in \mathcal{N}$ such that if $jk \in$, then either j=I or k=i. Individual i is the center of the star.

4.2.2. The Co-Author Model

In this model, the players, or the nodes are researches who write papers. Each node's productivity is a function of its links. A link represents collaboration between two researchers. Naturally, the amount of time a researcher spends on any given project is inversely related to the number of projects that researcher is involved in. As a result, in contrast to connections model, here indirect connections will enter the utility function in a negative way since they detract from one's co-author time.

The utility (or productivity) of player i given the network g is

$$u_i(g) = \sum_{j:ij \in g} w_i(n_i, j, n_j) - c(n_i)$$

where $w_i(n_i, j, n_j)$ is the utility derived by i from a direct contact with j when i and j are involved in n_i and n_j projects, respectively, and $c(n_i)$ is the cost to i of maintaining n_i links.

A more specific version of this model is stated as

$$u_i(g) = \sum_{j:ij \in g} [\frac{1}{n_i} + \frac{1}{n_j} + \frac{1}{n_i n_j}] = 1 + \left(1 + \frac{1}{n_i}\right) \sum_{j:ij \in g} \frac{1}{n_j}$$

and for $n_i = 0, u_i(g) = 0$. This form assumes that each researcher has a unit of time which they allocate equally across their projects. The output of each project depends on the total time invested in it by two collaborators, $1/n_i + 1/n_j$, and some synergy in the production process captured by the term $1/n_i n_j$. Here there are no direct costs of connection. The cost of connecting with a new author is that the new link decreases the strength of the interaction term with existing links.

Jackson and Wolinsky (1996) propose that in the co-author model:

(i) if N is even, the strongly efficient network is a graph consisting of N/2 separate pairs, and

(ii) (ii) a pairwise stable network can be partitioned into fully intraconnected components, each of which has a different number of members.

Stable networks will tend to be over-connected from an efficiency perspective. This happens because authors only partly consider the negative effect their new links have on the productivity of the links with existing co-authors.

4.2.3. Stability and efficiency

A network resulting from Jackson and Wolinsky's (1996) model is deemed to be stable if

(i) no individual agent has an incentive to sever a link and

(ii) no pair of agents have an incentive to form a new link.

Calvo-Armengol and Ilkiliç (2004) state that pairwise stable networks are robust to one-link deviations, where link severance is unilateral, while link creation is bilateral and under the mutual consent of two involved players. Among Dutta and Jackson (2001) pairwise stability is a very weak concept of stability because it only considers the addition or deletion of a single link at a time. Therefore, it is not a sufficient condition for stability but usually it is considered a necessary condition for stability.

4.3. Game theoretic network formation models

Within the field of game theoretic models the cooperative approach models were first developed, but approximately from the end of the 1990s the focus of the researchers has shifted on the non-cooperative models. The cooperative framework, especially games in characteristic form, did not prove entirely suited in games with externalities, i.e. virtually all games with genuine interaction among players (Marini 2007). This development is most probably due to the fact that the emphasis of research of networks has more or less shifted to more modern information transmission, buyer-seller and virtual networks such as Facebook and Twitter, for example, where there are externalities and costs are not divided equally. Non-cooperative models seem to be able to illustrate these kinds of networks where the cooperative approach fails to do so. Moreover, non-cooperative games are probably easier to analyze since many cooperative form games are reduced to non-cooperative form. I now present the noncooperative model of Bala and Goyal (2000).

Game theoretic models are based on the Nash equilibrium. In the Nash-equilibrium, no player has any incentive to change his or her own strategy unilaterally, i.e. to form new links or to sever old ones. This means that in equilibrium no player can gain by altering the strategy chosen (Myerson 1991). Dutta et al. (1997) define Nash equilibrium network so, that no agent wants to unilaterally break a link with any player. Because no player wants to break a link any cooperation structure can be sustained as Nash equilibrium. Consequently, no single player can form any new links through unilateral deviation, which among van den Nouweland (2004) motivates the use of Nash equilibrium for network formation models.

4.3.1. One and two-sided link formation

The model of Bala and Goyal (2000) was a pioneering one in the field of noncooperative game theoretic models[13]. This is mainly because their paper entails the dynamic approach which I present in the next chapter and they handle both one-sided and two-sided link formation. Many of the papers about network formation are concerned with the efficiency aspects of different network

[13] Although the model of one-sided and noncooperative link formation was introduced and some results on the static model were presented by Goyal (1993).

structures from the planner's point of view (e.g. Jackson and Wolinsky 1996) while Bala and Goyal consider the subject from the perspective of individual incentives. Their approach builds on the notion that social networks are formed by individual decisions that trade off the costs of forming and maintaining links against the potential rewards from doing so.

In Bala and Goyal's setting, each individual is a source of benefits that others can tap via the formation of costly pairwise links. Benefits are assumed to be nonrival[14]; a link with another agent allows access to the benefits available to the latter via his own links. Thus individual links generate externalities whose value depends on the level of decay and delay associated with indirect links. I will just present the simple model where neither decay nor delay are accounted for. The cost of link formation is incurred only by the person who initiates the link, unlike in the Jackson and Wolinsky (1996) setting where costs are assumed to be symmetric across the players. This unsymmetric or at the extreme one-sided division of costs makes the network formation process a noncooperative game.

A link between agent i and j can allow for either one-way (asymmetric) or the two-way (symmetric) flow of information. The former refers to a directed network whereas the latter to a nondirected network. The payoff is strictly increasing in the number of other people accessed (directly or indirectly) and strictly decreasing in the number of links formed. Each agent is assumed to possess some information of value to him and to other agents. He can augment this information by communicating with other people, which takes resources, time, and effort and is made possible via the setting up of pair-wise links.

4.3.1.1.One-way Flow model

Contexts where one individual may form a link with a second individual without the second individual's consent include sending a paper to another individual or web links. Such settings obviously lead to different incentives in the formation of networks, as mutual consent is not needed (Dutta and Jackson 2001). In the one-way flow model the strategy profile is $g = (g_1, \ldots, g_N)$ in $\mathcal{G}$ as a directed network. The link $g_{i,j} = 1$ is represented by an edge starting at j with the arrowhead pointing at i. $N^d(i; g) = \{k \in N | g_{i,k} = 1\}$ is defined as the set of agents with whom i maintains a link. There is a path from j to i in g either if $g_{i,j} = 1$or there exist distinct agents $j_1, \ldots, j_m$ different from i to j such that $g_{i,j_1} = g_{j_1,j_2} = \cdots = g_{j_m,j} = 1$. The notation "$j \overset{g}{\rightarrow} i$" indicates that there exists a

[14] For example, information sharing, concerning brands or products among consumers, the opportunities generated by having trade networks and the important advantages arising out of social favors are examples of nonrival benefits.

path from j to i in g. In addition, $N(i;g) = \{k \in N | k \to i\} \cup \{i\}$ is the set of all agents whose information i accesses either through a link or through a sequence of links.

In one-way flow, the payoff function to each agent is denoted as

$\Pi_i(g) = \Phi\left(\mu_i(g), \mu_i^d(g)\right)$ which can be put as $\Pi_i(g) = \mu_i(g) - \mu_i^d(g)c$.

here $\mu_i(g)$ is the number of agents observed by the agent i which can thus be interpreted as the benefit to agent i. Notion $\mu_i^d(g)$ is the number of agents with whom i has formed links and measures the cost associated with maintaining them.

4.3.1.2. Two-way flow model

A nondirected network is denoted $\overline{g} = cl(g)$, and defined by $\overline{g}_{i,j,} = \max(g_{i,j}, g_{j,i})$ for each i and j in N. There is a two-way path in g between i and j if either $\overline{g}_{i,j,} = 1$ or there exist agents $j_1, \ldots, j_m$ distinct from each other and i and j such that $\overline{g}_{i,j_1} = \cdots = \overline{g}_{j_m,j} = 1$. The notion $i \overset{g}{\leftrightarrow} j$ indicates a two-way path between i and j in g. $N^d(i;g)$is defined as in one-way model. The payoffs are

$\overline{\Pi}_i(g) = \Phi\left(\mu_i(\overline{g}), \mu_i^d(g)\right)$ and $\overline{\Pi}_i(g) = \mu_i(\overline{g}) - \mu_i^d(g)c$ where upper bar denoted the nondirected network.

This definition of payoffs means that while one agent bears the cost of a link, both agents obtain the benefits associated with it. Examples of two-way information flow with asymmetric cost division include phone call.

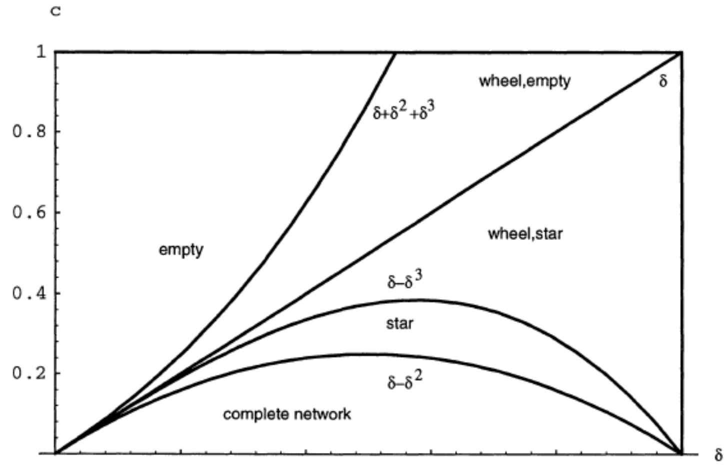
4.3.1.3. Emerging networks

Simply put, agent i's payoff is the number of agents he observes less the total cost of link formation. The payoff function implicitly assumes that the value of information does not depend upon the number of individuals through which it has passed, that is, there is no information decay or delay in transmission[15]. The payoff is strictly increasing in the number of other people accessed and strictly decreasing in the number of links formed. If $c \in (0,1)$, then agent i is willing to form a link with agent j for the sake of j's information alone. When $c \in (1, n-1)$ agent i requires j to observe some additional agents to induce him to form a link with j. When $c > n-1$, the cost of link formation exceeds the total benefit of information available; here, the dominant strategy for i is not to form any links at all.

[15] See Bala and Goyal (2000) p. 1210 for details about he models with decay and delay.

Bala and Goyal (2000) find that Nash networks are either minimally connected or empty. In the one-way flow model, the only strict Nash architectures are the wheel and the empty network[16]. A strict Nash network is one where each agent gets a strictly higher payoff with his current strategy than he would with any other strategy.

Given the payoff function, the wheel is the unique strict Nash if $\Phi(\hat{x}+1,\hat{x}) > \Phi(1,0)$ *for some* $\hat{x} \in \{1,\ldots,n-1\}$. *If* $\Phi(x+1,x) < \Phi(1,0)$ *for all* $x \in \{1,\ldots,n-1\}$ *and* $\Phi(n,1) > \Phi(1,0)$ *then empty network and the wheel are both strict Nash. If* $\Phi(x+1,x) < \Phi(1,0)$ *holds for all* $x \in \{1,\ldots,n-1\}$ *and* $\Phi(n,1) < \Phi(1,0)$ *then the empty network is the unique strict Nash.*



Picture 7. Strict Nash networks, one-way model (n=4). (Bala and Goyal 2000)

This result underlines a general property of Nash network when agents are symmetrically positioned in relation to information and costs of access: in equilibrium, either there is no social communication or every agent has access to all the information in the society. If the network is not minimally connected, then some agent could delete a link and still observe all agents, which would contradict Nash.

In the two-way model, the asymmetry in payoffs referred to earlier is relevant to defining the architecture of the network. Here, there are three types of star networks, depending upon which agents bear the costs of the links in the network. The three types of star networks are center-sponsored, periphery-sponsored and a mixed type star. Also, there can be a large number of Nash equilibria. However, strict Nash network architectures are either center sponsored star and the empty network.

A center sponsored star is strict Nash if and only if $\Phi(n,n-1) > \Phi(x+1,x)$ *for all* $x \in \{0,\ldots,n-2\}$. *The empty network is strict Nash if and only if* $\Phi(1,0) > \Phi(x+1,x)$ *for all* $x \in \{1,\ldots,n-1\}$.

Some critique towards the model has also been suggested: the fact that no consent for link formation is required is unacceptable for social networks. Also, applications are reputedly very

[16] A wheel network (see Picture 4) is such where the agents are arranged as $\{i_1,\ldots,i_N\}$ with $g_{i_2,i_1} = \cdots = g_{i_n,i_{n-1}} = g_{i_1,i_n} = 1$ and there are no other links. Bala and Goyal denote the wheel network as g^w. A star network has a central agent i such that $g_{i,j} = g_{j,i} = 1$ for all $j \in N\backslash\{i\}$ and no other links.

limited; links have to be viewed as the confirmations of existing relationships, e.g. one of the agents already linked to each other makes a phone call and bears the costs. Furthermore, the payoff functions investigated are rather unrealistic. Yet, the one-sided link formation could describe the way in which part of the virtual links in social networking sites are formed. For example, people in Twitter might follow tweets of some particular user without sender's consent if s/he has allowed open access. Here, the link is obviously one-sided.

4.3.2. Stability and efficiency

In Bala and Goyal's model a complete graph is never efficient or stable due to the noncooperative approach. Let g_{-i} denote the network obtained when all of agent i's links are removed. The network g can be written as $g = g_i \oplus g_{-i}$ where the $\oplus$ indicates that g is formed as the union of the links in g_i and g_{-i} . The set of agent i's best responses to g_{-i} is denoted $BR_i(g_{-i})$. Furthermore, a network $g = (g_1, \dots, g_n)$ is a Nash network if $g_i \in BR_i(g_{-i})$ for each i.

The welfare measure W(g) is defined in terms of the sum of payoffs of all agents. A network is efficient if $W(g) \geq W(g^{'})$ for all $g' \in \mathcal{G}$. The corresponding welfare function for two-way communication is denoted $\overline{W}$. For the linear payoffs, an efficient network is one that maximizes the total value of information made available to the agents, less the aggregate cost of communication.

Bala and Goyal find that if the cost of forming links is low or if the network is highly reliable then there is no conflict between efficiency and stability (Dutta and Jackson 2001).With the two-way flows, the subject of efficiency is quite complex. For example, a center-sponsored star can have a different level of welfare than a periphery-sponsored one, since the number of links maintained by each agent is different in the two networks. However, given the linear payoffs, it can be shown that if $c \leq n$ a network is efficient if and only is it is minimally two-way connectected (a star is efficient) while if $c > n$ then the empty network is uniquely efficient (Bala and Goyal 2000).

4.4. Discussion about the different equilibrium concepts

Pairwise stability and Nash equilibrium have very dissimilar approaches. Pairwise stability applies directly to the network and players' payoffs from networks. However, in most of the game theoretic models, such as in a noncooperative linking game (e.g. Myerson 1991) agents independently announce which bilateral links they would like to form and then standard game-theoretic equilibrium concepts can be used to make predictions about which networks will form. Many examples show that the set of the Nash equilibrium outcomes of the linking game can be completely different from the set of pairwise stable networks (Bloch and Jackson 2006). In addition, the connections between various solution concepts are rarely explicitly stated or studied (Bloch and Jackson 2006).

The concept of Nash equilibrium is perhaps most commonly applied to measure the stability in strategic form models of network formation albeit it might be that pairwise stability is more intuitive and suits better to describe the stability of social networks. Both notions of stability do have their strengths and weaknesses.

Pairwise stablility is not based on an explicit noncooperative game of network formation. Instead, it is a direct stability check which rules out networks which can intuitively be considered as unstable. Pairwise stability is a simple concept to use in applications because it has good computational properties. The fact that it only considers very simple deviations is its main shortcoming, and a network may be classified as stable too easily. Central to the notion of pairwise stability is the fact that a deviation can include two players who come together to form a new link. The concept of Nash equilibrium does not allow for such consideration but it sees players as individuals who do not cooperate, this thus holds for noncooperative models (Jackson and Wolinsky 1996). Nevertheless, as Dutta and Jackson (2001) and the authors of the paper themselves have criticized the model, the basic stability notion of Jackson and Wolinsky (1996) requires only that a network is immune to one deviating action at a time. That is, it is not required that a network be immune to more complicated deviations, such as a simultaneous severance of some existing links and introduction of a new link by two players. It is not required either that a network is unaffected to deviations by more than two players simultaneously. In fact, Jackson and Wolinsky admit that the notion of pairwise stability does not even contemplate the severance of more than one link by a single player. The consequence of strengthening the stability notion so to allow for more complex deviations is the fact that in some cases there may not exist any stable network.

Unlike the pairwise stability notion, the concept of Nash equilibrium fails to account for the fact that value of a network can depend on how the players in a network are interconnected. This is basically implicit in the Jackson and Wolinsky model but to my knowledge this has not been incorporated to the game theoretic models. Considering how the value depends on the pattern of connectedness might have substantial effect which networks are considered as efficient. Jackson (2008a) argues that the noncooperative game Nash equilibrium fails to model the network formation because it does not take cooperation and communication between the players into consideration. Usually, the network formation games may have too many Nash equilibria of which many are easily seen as unreasonable, such as empty networks, regardless of the networks. In other words, Nash equilibrium allows players to refuse to form links and thus effectively delete links, but as Jackson (2008b) argues that it does not capture the fact that it may be mutually advantageous to form a relationship.

To address the fact that the consent of both players is needed to form a link, that an undirected network do not account for, one should consider coordinated actions on the part of pairs of players. This is the motivation behind pairwise stability (Jackson 2008b). To be able to explore changes beyond one link at a time, one needs to come up with a revision of a stability concept. One useful refinement of both the set of pairwise stable networks and Nash equilibrium of the linking game called pairwise Nash equilibrium is exactly the intersection of the set of the Nash equilibrium outcomes of the linking game and the set of pairwise stable networks, when such networks exist (Bloch and Jackson 2006). Pairwise Nash stability thus combines pairwise stability and Nash equilibrium. A rational network is one that is both Nash stable and pairwise stable, i.e. a network is pairwise Nash stable if it is both Nash and pairwise stable (Calvo-Armengol and Ilkiliç 2004).

In the next chapter, I will develop the fundamental models further and discuss some extensions of the strategic form models that introduce dynamics, heterogeneity across players and endogenous link strength.

5. Other game theoretic models of network formation

Thus far, the strategic form networks have demonstrated static features. In this part, I discuss models introducing dynamics, heterogeneous players and models where link strength is defined endogenously. None of these modifications is present in labor market models with the network approach as such but the assumptions in the extensions in this chapter are more similar to the labor market models and above all, more realistic than the ones in the basic Jackson and Wolinsky and Bala and Goyal model. The lack of these assumptions might be due to the fact that Calvó-Armengol (2004) seems to be the only one who has even succeeded in creating an endogenous network formation model in labor labor information networks. Yet, all the labor market models have several – or at least two – periods, thus understanding dynamics in network formation is useful. Also, unlike many basic labor market theories, such as search theory, the network approach assumes heterogeneous employees; that is why I speak about applications with heterogeneous players. The lack of complicated models can be explained by a rather short lifespan of this field of research, or because simpler models are able to adequately explain the phenomena in question. It is most probably true that the more complex model the less explanatory power it has. Nevertheless, I believe that the assumptions of heterogeneous agents and endogenous link strength should be included in the endogenous network formation model.

5.1. Dynamic models

By understanding the dynamic process of network formation one might be able to better predict the impact of certain acts. The economy – i.e. economic actors - is not in a static state but the structure evolves and changes over the time (Easley and Kleinberg 2009); agents are assumed to be farsighted. Also, most of the labor market models with the network approach have two or more periods.

The study of dynamics in network formation is the second contribution among one-sided link formation in Bala and Goyal's (2000) paper. Other papers discussing dynamic models include that of König, Tessone and Zenou (2009)[17]. The research of dynamics is justified by the fact that they believe that dynamics may help select among different equilibria of the static games. Bala and Goyal use best-response dynamics to study the issue; an individual chooses a set of links that maximizes his payoffs given the network of the previous period. The network formation game is assumed to be repeated in each period t=1,2,...In the model there is some fixed probability $r_i \in (0,1)$ agent i to

[17] In the model of König et al. (2009) links are formed on the basis of agents' centrality (Bonacich centrality) while the network is exposed to a volatile environment introducing interruptions in the connections between agents. They show that there exists a unique stationary network whose topological properties match features exhibited by real-world networks. There exists a sharp transition in efficiency and network density from highly centralized to decentralized networks.

exhibit inertia, i.e. individual chooses the same strategy as in the previous period. Inertia is not expressed with probability $p_i = 1 - r_i$; then an agent chooses a myopic pure strategy best response to the strategy of all other agents in the previous period. Also, if more than one strategy is optimal, an agent randomizes across the optimal strategies.

Formally, the dynamic model is presented as follows. For a given set A, $\Delta(A)$ denotes the set of probability distributions on A. For each agent i there exists a number $p_i \in (0,1)$ and a function $\phi_i : \mathcal{G} \rightarrow \Delta(\mathcal{G}_i)$ where ϕ_i satisfies, for all $g = g_i \oplus g_{-i} \in \mathcal{G}$:

$$\phi_i(g) \in \text{Interior } \Delta(BR_i(g_{-i})),$$

For $\hat{g}_i$ in the support of $\phi_i(g)$ the notation $\phi_i(g)(\hat{g}_i)$ denotes the probability assigned to $\hat{g}_i$ by the probability measure $\phi_i(g)$. If the network at time $t \geq 1$ is $g^t = g_i^t \oplus g_{-i}^t$, the strategy of agent i at time t+1 is assumed to be given by

$$g_{-i}^t \begin{cases} \hat{g}_i \in support\ \phi_i(g) & with\ probability\ p_i \times \phi_i(g)(\hat{g}_i) \\ g_i^t & with\ probability\ 1 - p_i \end{cases}$$

The equation states that with probability $p_i \in (0,1)$ agent i chooses a naive best response to the strategies of other agents. The function $\phi_i$defines how agent i randomizes between best responses if more than one exists. Additionally, with probability $1 - p_i$agent i maintains his previous strategy. Bala and Goyal assume that the choice of inertia as well as the randomization over best responses by different agents is independent across agents. The dynamic process in the two-way model is the same except that the best- response mapping is denoted $\overline{BR}_i(\cdot)$ instead of $BR_i(\cdot)$.

Bala and Goyal (2000) suggest that the dynamic process converges to a limit network. In the one-way flow model, for any number of agents, the dynamic process converges to a wheel or to the empty network, with probability 1. In the two-way flow model, the dynamic process converges to a center-sponsored star or to the empty network with probability 1.

5.2. Endogenous link strength and heterogeneity across agents

Nearly all formal models of networks treats links as binary quantities, i.e. they are either present or absent, which is clearly a simplification of the networks observed in reality. Relationships are characterized not only by the presence or absence of a link but also by the intensity, frequency or reliability whereby they occur (Rogers 2005). Bloch and Dutta (2008) analyze the link formation in communication networks when players choose how much to invest in each relationship, where the quality of links is endogenously chosen by the agents. Most of the models assume the strength of the link to be exogenously given. The paper of Bloch and Dutta is one of the first to address to the strength of links. This issue has been studied in the field of sociology earlier; Granovetter (1973) has examined the strength of the weak ties, links, and one of his conclusions is that job information is more likely to transmit through weak than strong links. Rogers (2005) introduces heterogeneity across agents because the utility of links depends not only on the amount of links an agent has but also who the acquaintances are.

5.2.1. Link strength as an endogenous variable

The fact that Bloch and Dutta consider link strength as endogenously chosen by the players can make this model useful starting point from the labor market perspective –given that the strength of the link is a significant factor in getting job information. Intuitively thought, the strength of a link should have a considerable effect on the quality of the information and the speed at which it transmits. The impact could be somewhat the same as in the case of homophilous networks. That is, the information between same-type agents is more likely to be correlated and therefore the information flow between different-type agents ought to be more valuable. Here, players perhaps choose to form strong links with agents with similar characteristics and weak links with players that distinct from themselves. Therefore, the information such as job information transmitted through weak links might be more valuable in comparison with strong links. In other words, the players who are connected with a strong link are more likely to possess similar information and thus, also less new job information tends to spread through strong links. Then again, the flow of information between weak links might be slower compared to that of strong links.

The strength of a social link in social networks depends on the frequency and the length of social interactions. The model is a generalization of Bala and Goyal's (2000) one-sided, two-way flow model of link formation (later they also cover the Jackson and Wolinsky 1996 setting). Each individual has a total resource (time, money) X > 0 and has to decide on how to allocate it in establishing links with others. The assumption of agents facing a fixed cost in the formation of links leads Bloch and Dutta (2009) to specify the strength of a link as an additively separable and convex function of agents' investments.

Let x_i^j denote the amount of resource invested by player i in the relationship with j. The strength of the relationship is then denoted

$s_{ij} = \phi(x_i^j) + \phi(x_j^i)$

where $\phi(.)$ is a nondecreasing, convex function. The separability of the function implies that an agent's decision to allocate his endowment over direct links is independent of his neighbors' decisions; however, these choices affect the value of indirect links so agent's investment strategy is not independent of the choices of other agents.

Also Bloch and Dutta themselves admit that one obvious drawback in their analysis is the fact that agents are homogenous. This assumption leads to the conclusion that links will all be of the same quality (with linear investments) or some agents will be better connected than others due to the hub's investments (convex investments). Neither the distributions of link intensities do justice to the broad array of social networks observed in reality, and as such do not demonstrate the idea of strong and weak links. Rogers (2005) introduces heterogeneity across players.

5.2.2. Heterogeneous agents – asking and giving model

Exploring models with heterogeneous agents is relevant in the labor information networks. For instance, some people are more likely to pass on job offers because of their professional contacts (Rogers 2005). In other words, some aspects of what a person has to offer depend on who their acquaintances are and not just on the amount of links a person has. Therefore, Rogers (2005) introduces heterogeneity across agents' intrinsic qualities and across linking costs. Asymmetries between players are natural; some players may be more productive or more informed compared with others. In the search theoretic labor market models the players are typically homogenous whereas the labor market theory with network approach builds to heterogeneous agents; there are both low-skilled and high-skilled workers. In addition to Rogers, Galeotti and Goyal (2002) discuss a model with heterogeneous agents. Their model is very much like the one of Rogers (2005) except the flow of information is two way and link strength is given.

In the asking and giving model Rogers (2005) combines endogenous link strength with heterogeneous players; the characteristics of agents are allowed to vary along two dimensions – intrinsic qualities and linking costs. The model is a one-way flow model like Bala and Goyal's (2000) with the exception that Rogers separates the flow of benefits into asking and giving models, that is, the one-way flow of benefits or information can be in both directions but not at the same time. This pattern of diffusion could fit quite well to the way job information is transmitted; when an employed agent decides to give job information, strictly speaking only the receiving agent benefits (omitting the possible benefit or gratification the giving agent receives), thus the flow of benefits is not two way.

The asking model is such that an unemployed agent asks about vacant jobs from his or her connections. In this model, agents are endowed with publicly observed intrinsic qualities, or value measured in terms of quality of information they possess $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}_+^n$ and linking budgets $\beta = (\beta_1, \dots, \beta_n)$ with $0 < \beta_i < 1$ for each $i \in N$, which are the sources of heterogeneity across agents. The difference in linking budgets stems from the fact that some agents have more social interest or ability, and thus naturally form stronger links to others. In the asking model, given the linking choices and intrinsic values of others, agent i chooses the relative amounts of benefits to receive from others from her budget of resources. In the giving model agent chooses how to give benefit given the linking choices of all others. The benefit from a link thus depends on the value of the agent linked to and the strength of the link. Additionally, the value of the agent depends on its set of links and the values of its neighbors, and so forth[18].

These modifications give new insights into the question of efficiency and stability. In the giving model there are differences between Nash and efficient networks which arises under heterogeneous budgets because for efficiency, linking choices are determined by total connectivities, whereas the equilibrium choices depend on individual connectivities. The discrepancy between equilibrium and efficiency is not present in the asking model, however. Also, in the giving model the linking choice of individuals is independent of the intrinsic values of other agents only the strength of the paths seems to matter. Neither asking nor giving model produces networks with wheel structures in equilibrium. A star network, on the other hand, is always efficient in the asking model when it forms. In the giving model star is never efficient because stars only form when budgets are not homogeneous; in this model, the empty network is Nash for budgets of every size. The equilibrium networks between the two models differ in that Galeotti and Goyal (2002) model's strict Nash network is a minimal network where every component with three or more players is a center-sponsored star.

Center-sponsored star architecture plays a prominent role even in the presence of heterogeneous values and differences in the cost of forming links across players (Galeotti and Goyal 2002). The star network is indicative of the role of the "hubs" in real life networks, where a few nodes have many connections. The occurrence of hubs has important implications for network performance, mainly due to their role in decreasing the distances between nodes in the network (Rogers 2005).

[18] Formally, let v_i denote the (endogenously) determined component of i's utility derived from her relations in the network $f(\Phi)$ so that I's total utility is $u_i = \alpha_i + v_i$. In model asking, $v_i = \sum_j (\alpha_j + v_j) f(\phi_{ij})$, whereas in model giving $v_i = \sum_j (\alpha_j + v_j) f(\phi_{ji})$. Here, $f(\phi_{ij})$ denotes the intensity of the directed link ij.

5.3. Discussion

I have tried to concentrate on the models that are relevant from the labor market and information flow perspective, trying to identify factors that are the most important in the formation of social networks that should be included in the models. Obviously, there is a large body of work in economics which examines the game theoretic modelling of network formation with various assumptions and contexts. Presenting even all the types of the extensions is beyond the scope of this work. Of the models presented above the dynamic formation model is perhaps the most central because agents are farsighted and optimize beyond one period. It is important to see to which structure the network convergences because it essentially affects labor market dynamics; static models fail to consider this aspect altogether.

To conclude, I name a few examples of the models that I consider important which were not discussed above. One class of applications is such that introduces link formation with costs, e.g. Slikker and van den Nouweland (2001) though applying only to cooperative form models. Networks also enable the trading of many goods, i.e. the trading happens in networks of buyers, seller and some third party actors, and these trading networks are thus important for the overall economic outcome. Kranton and Minehart's (2000) model of buyer-seller networks is perhaps the most prominent in the literature. One typical modification of the basic game theoretic model is link formation with transfers or side payments, e.g. Bloch and Jackson 2006, which essentially stand for the same thing. I believe that when links in a social networks form, no side payments are made; side payments could be the means when a social planner wishes for a certain network structure to emerge. The models that I have presented here all assume perfect information, meaning that individuals are aware of the structure of the network. Assuming perfect information in linking decisions might not be feasible in all cases, one valuable modification of Galeotti et al. (2009) introduce network formation game with imperfect information where the level of information shapes individual behavior and payoffs.

I have now discussed network formation and presented several models of network formation with different assumptions. In the next part, I turn to analyze some effects that networks have on the economy. In other words, I will deliberate how one can explain certain phenomena of the networked labor markets with network approach. Here, I will compare theories with the network approach to other widely applied economic theories.

6. Networks and labour markets

Understanding the role of networks in various settings is very central; a particularly relevant topic in the economic research is the study of the effect of social networks in in labor markets. Job markets are typically characterized by asymmetric information on job opportunities and workers' ability (Goyal 2007). Having more or better connections may allow individuals to receive more information on vacancies since access to information is heavily influenced by social structure (Ioannides and Loury 2004). The network pattern of relationships is important as it describes who passes information to whom, which is ultimately crucial in determining a person's long-term employment prospects (Calvó-Armengol and Jackson 2004). During last decade, these issues have been rather widely recognized in the literature. In essence, the topics "volatility" of unemployment and inequality measured in terms of wage and dropout rate differentials in the labor market are for the most part the motivation and the main focus of this chapter.

In this part, I also compare the explicative power of search theoretic and network approach regarding the trends prevailing in the labor markets. In other words, my aim is to show if network approach can possibly provide a better theory for the persistence of unemployment and wage differences between groups. Overall, the network approach does not strive for undermining search or matching model but it offers and additional explanation to phenomena that search approach is thought not to be able to account for. Here I presume that network approach provides a more valid theory to the phenomena in question compared with search theory. In effect, the papers with network approach show how network considerations imply different outcomes than the simple one-to-one search models that typify most economic analyses of job acquisition. Still, it remains unclear whether e.g. duration dependence is due to unobserved heterogeneity or the network structure of the economy. For the time being, the position of the network approach in the field of research is validated because it offers at least a complimentary explanation to these questions.

I first introduce the search and matching model to provide the basic understanding about labor market models. Thereafter I discuss labor market models with the network approach and about the fact how certain features of labor markets are explained by the network approach. Now, I shortly discuss the characteristics of networked labor markets.

6.1. Characteristics of networked labor markets

During the last few decades, there have been many developments in the field of labor economics. The means how labor economics seeks to understand the functioning and dynamics of labor markets has shifted from neoclassical supply and demand models towards search and matching theory and beyond. Yet, the search and matching framework has been criticized for being unable to match key labor market statistics, namely the volatility of unemployment (variation in unemployment duration) and job vacancies. This casts doubt on the viability of the search and matching model to provide a theory for labor market dynamics (Lubik 2009). Now there is a move towards labor market models with the network approach because standard economic analysis does not account for the role of social networks in the employment process and is hence unable to explain certain persisting phenomena in labor markets. This view states that informal hiring channels such as job search through networks are important for both workers and employers and consequently, the pattern of social ties between individuals may play an important role in determining the labor market outcomes (Montgomery 1991).

Labor markets are characterized e.g. by wage and employment status differentials between various groups of individuals, between men and women, different races and ethnic groups, the older and younger and people with a different educational background. Among Stovel and Fountain (2006) the classical explanations for segregation in labor markets in the demand side are employer preferences for-or against particular groups of workers (blacks, women, immigrants) and in supply side difference in skill level between groups and difference in preferences to workers for types of jobs. If this was the sole reason, there would still be to explain why the characteristics of different groups should differ. However, which people are hired into which jobs does not depend only on individual and human capital characteristics of workers or the preferences of employers; it is also a function of the complex process by which people and jobs hear about and are matched with one another (Granovetter 1981).

Source of Hire	Source Value Index
Employee Referrals	3.56
Social Networking	1.58
Niche Job Boards	0.82
Commercial Resume Database	0.80
General Job Boards	0.59
Newspapers	0.57
Career Fairs	0.56
Search Firms	0.53

(Source Value Index = Percent of Hires / Percent of Spend)

Source: 2007 Recruiting Trends Survey, Sponsored by the DirectEmployers Association™

Table 2. Networks as a source of hire.

Thus, social networks are an important means in acquiring a job; Montgomery (1991) states that approximately 50 percent of all workers currently employed found their jobs through friends and relatives. From the firm perspective, employee referrals are a useful device for screening job applicants. Referrals are more valuable than anonymous matching because in conveying information about employees' qualifications, they reduce recruiting costs and training costs, and lower monitoring costs. The use of friends and relatives to search for jobs often varies by location and by demographic characteristics, but generally this kind of job search is quite productive (Ioannides and Loury 2004).

6.2. Labor market models

The fundamental difference between search theory and network approach in describing labor markets is most likely the fact that search theory formally models *frictions* associated with job seekers' access to information about the availability of the jobs of different types and about the conditions of employment (Ioannides and Loury 2004) whereas the network approach acknowledges the fact that access to job information is heavily influenced by social structure. However, quite often in economic theory it fails to be well-grounded under what circumstances networks are likely to have their largest effect on labor market outcomes, and when the effect is positive or negative because networks' influence mechanism on economy is still under study and debate.

6.2.1. Introduction to search and matching models

Although the simple static neoclassical labor market model[19] is useful in illustrating the operations of labor markets in the basic microeconomic level, it is deficient in explaining macroeconomic events, e.g. in equilibrium, there is no unemployment. The new "economics of imperfect information" was invented to deal with uncertainty and matching; this research has led to a growing literature under the heading "job search theory". (Devine and Kiefer 1991)

The search approach was developed as a static model for the behavior of unemployed workers (Stigler 1961), yet Mortensen (1970) was among the first to introduce a nonstatic search model. The model also extended to include on-the-job search. Formally, the basic model among Devine and Kiefer is as follows (see e.g. Rogerson et al. (2005) for a slightly different approach). A worker seeks to maximize the expected value of income discounted with r. The income flow while unemployed is b and it is constant over the duration of a given spell. Offers are received while unemployed according to a Poisson process with parameter δ, which is the arrival rate of offers. A job offer is summarized by a wage rate w. Successive job offers received over the course of a spell of unemployment are independent realizations from a known wage offer distribution with finite mean and variance. When accepted, a job will last forever which is not a crucial assumption, however. The simplest way of generating transitions from employment into unemployment is to assume that jobs end for some exogenous reason (Rogerson et al. 2005). The search strategy does not depend on the time unemployed. This model assumes homogenous firms and workers but there are applications that introduce heterogeneous workers and firms although in the minority. When it comes to matching models, the general idea is that both workers and firms are heterogeneous and some matches

[19] In neoclassical model workers choose their labor supply by maximizing utility, which is a function of income and leisure, subject to a budget constraint in which prizes are given. Aggregate demand and aggregate supply in a labor market are derived by a summation of the individual demand and supply functions. (New Palgrave Dictionary of Economics)

between workers and firms are more productive than others. Before a match is formed, all workers look alike to a firm and all firms look alike to workers; the true time invariant productivity of a worker in a particular match can be learned only over time thus there exists imperfect information. (Devine and Kiefer 1991).

6.2.2. The network approach

Using networks, informal channels that is, to acquire a job is less expensive for both job seekers and firms than formal methods; it also ameliorates the problem of adverse selection because workers tend to refer others who are similar to themselves. This is an evidence of homophily – links tend to from between individuals who are similar to each other. Employers will thus solicit referrals from high-ability employees.

Due to the unclear results of how networks affect the labor market outcome and when the effect is at the largest, much of the recent research about job networks points out four important considerations to take into account in the analysis; employer, relational, contact and worker heterogeneity (Ioannides and Loury 2004). Employer heterogeneity determines which search methods –formal or informal – an employer prefers, relational and contact heterogeneity describes the variations in the social relationship endowments of individuals, i.e. the position of an individual in a network, the network density etc., and finally, worker heterogeneity indicates what kind of differences in worker productivity or characteristics there are.

There are mainly two kinds of models, one where the network structure is taken as given and the other where the structure is allowed to vary. In the former, the structure remains unchanged and in the latter network formation models are introduced. However, compared with the models presented in the previous chapters, the formation models are generally simpler and impose stricter assumptions. Yet the hypothesis of static, fixed, network structure is not very plausible because the structure evolves over time. Due to new social networking sites that provide a means to link creation, such as Facebook and LinkedIn in professional contacts, network structure might change probably faster than in the case of physical networks, not at least due to lower linkage costs.

6.2.2.1. Fixed network structure

In the following two models the network structure is assumed static, thus it does not change within the periods. The model assuming fixed network structure illustrates the basic logic of labor market models incorporating social network structure. Montgomery (1991) tries to embed social structure in a model of the labor market which offers a framework for exploring the equilibrium relationship among social structure, wages and profits. This model builds to the observation that workers tend to refer others like themselves. Calvó-Armengol and Jackson (2004) extend their model to several periods and in addition introduce an exogenous probability of people losing their jobs.

In the model of Montgomery, there are two periods, and each worker lives one period. Half or workers are of high and half of low ability. High-ability workers produce one unit of output while low-ability workers produce zero units. Workers are observationally equivalent, meaning that employers are uncertain of the ability of any particular worker thus there is information asymmetry present. Each firm may employ at most one worker, and firm's profit is equal to the productivity of its employee minus the wage paid. Each firm must also set wages before learning the productivity of its worker and any output-contingent contracts are prohibited. If fully output-contingent pay scheme was allowed and possible then firms would not have any incentive to hire through referrals. Each period-1 worker knows at most one period-2 worker with probability $\tau \in [0,1]$. For each period-1 worker holding a tie, the specific period-2 individual known is selected stochastically through a two-stage process. In the first stage, the period-2 worker's type is chosen. Conditional upon holding a tie, period-1 worker knows a period-2 worker of his own type with probability $\alpha > \frac{1}{2}$. In the second stage, the specific period-2 worker is chosen randomly from those of the appropriate type. The social structure is characterized by two parameters; network density (τ) and inbreeding bias (α). Thus, some workers are well connected while others are not. In equilibrium a firm will attempt to hire through referral if and only if it employs a high-ability worker in period 1.

Calvó-Armengol and Jackson (2004) and Montgomery (1991) models differ in that Montgomery does not consider any arrival rate of job-offers or referrals or break-up rate, maybe because he only has two periods in the model. This model may still provide useful insight in representing how exactly wages should react to the use of informal channels in job search when also the outside option, i.e. getting a job from a formal channel, is available. Calvó-Armengol and Jackson (2004) also assume for simplicity that all jobs are identical. The n agents are heterogeneous as in Montgomery's model. The vector s_t describes the employment status of the agents at time t. Information about job openings arrives; any given agent hears about a job with probability $a \in (0,1)$. If an agent is unemployed he or she will take the job but if an agent is already employed s/he will pass the information along randomly to a friend, relative or acquaintance who is unemployed. If all of an agent's acquaintances are already employed, then the job information is simply lost. The network is assumed to be nondirected $g_{ij} = g_{ji}$. The probability of the joint event that agent i learns about a job and this job ends up in agent j's hands is described by $p_{ij}\,(s)$, where

$$p_{ij}\,(s) = \begin{cases} a & if\ s_i = 0\ and\ i = j \\ \frac{a}{\sum_{k;s_k=0} g_{ik}} & if\ s_i = 1, s_j = 0, and\ g_{ij} = 1; and \\ 0 & otherwise \end{cases}$$

where the vector **s** describes the employment status of all the agents at the beginning of the period. Indirect relationships also play a role. In the short run they are competitors for job information but in the long run they help keep an agent's friends employed which is a benefit. Finally, the model introduces an exogenous breakup rate, b, between 0 and 1, the probability that any given employed agent will lose his or her job at the end of a given period. In this model, as time unfolds, employment evolves as a function of both past employment status and the network of connections. The probability of unemployment can fall with the number of links but the average unemployment is shown to increase with closed-knittedness, reflecting the fact that the wider the breadth of current social ties, the more diversified are the sources of information. Usually, direct connections are more important than indirect ones, but in situations with lower arrival rates and higher break-up rates, indirect connections can become more important. So this is where the strength of weak links might show. Finally, Calvó-Armengol and Jackson state that the long run positive correlation of any interconnected agents' employment implies that there is a clustering of agents by employment status, and employed workers tend to be connected with employed workers and vice versa. A more general model with fixed network structure is available from the same authors (Calvó-Armengol and Jackson 2007). Other models with fixed network structure include e.g. Montgomery (1994) where social structure consists of groups of two connected individuals, dyads, and there are both strong and weak ties.

Next I show a more complicated model with variable network structure by Calvó-Armengol (2004).

6.2.2.2. Variable network structure

Permitting for network structure to alter makes the analysis of networks' impact on the labor market outcome less straightforward but brings some important insights into consideration. Assuming static network structure does not allow for considering e.g. the dynamics of unemployment by means of the network approach, since two networks with the same total number of links but different geometry may induce different aggregate unemployment levels (Calvó-Armengol 2004). To illustrate the idea of network formation models in labor market settings, I present the Calvó-Armengol (2004) model which is among the few ones[20] that have successfully employed a model of strategic network formation to the job market context (Ioannides and Loury 2004).

The basic notions apply for the contact networks in Calvó-Armengol (2004). Here, the same principles in job transmission as well as breakdown and arrival probabilities apply as in Calvó-Armengol and Jackson (2004). Initially, all players are employed. Let $\alpha = a(1-b)$ and $\beta = b(1-a)$. An employed player has an extra job slot for his contacts with probability α, and an unemployed player need his contacts to find a job with probability β. The probability that i gets a job through contacts is $P_i(g) = 1 - \prod_{j \in N_i(g)} q(n_j(g))$ where $q\left(n_j(g)\right) = 1 - \alpha \frac{1-(1-b)^{n_j(g)}}{bn_j(g)}$ is the probability that i does not get a job from $j \in N_i(g)$. Calvó-Armengol finds contrary to Calvó-Armengol and Jackson (2004) that direct contacts increase individual employment prospects. Two-links away contacts are competitors for information and harm employment prospects. The aggregate unemployment rate u(g)[21] is determined by $\left(P_1(g), \ldots, P_n(g)\right)$ because it also determines how job information flows through personal contacts.

The link formation process proceeds as in Bala and Goyal (2000) so that players individually announce all the links they wish to form but with the exception that mutual consent is needed. So each link ij results in a cost c>0 to both i and j, equal across players.The expected payoff of a player i is

$$Y_i(g) = \underbrace{(1-b)}_{i\ keeps\ job} + b\underbrace{\left[\underbrace{a}_{market} + \underbrace{(1-a)Pi(g)}_{contact}\right]}_{i\ fired\ and\ reemployed} - cn_i(g)$$

Here, it can be seen again that Nash equilibrium is too weak an equilibrium concept because the empty network is always a Nash equilibrium. Calvó-Armengol refines the equilibrium concept to build upon the pairwise stability concept. $g \in \mathcal{G}$ is a pairwise-equilibrium network if and only if there is a Nash equilibrium strategy profile which supports g and, for all $ij \neq g$, $Y_i(g+ij) > Y_i(g)$ implies $Y_j(g) > Y_j(g+ij)$. He concludes that pairwise-equilibrium networks always exist. If $c > \alpha\beta$ the empty graph is the only pairwise-equilibrium network. Equilibrium networks are non-empty if the net value of a first link, $\alpha\beta - c$ is nonnegative. As far as efficiency is considered, Calvó-Armengol concludes that pairwise-equilibrium networks can generally be inefficient. First, individual incentives to form contacts may sometimes be excessive in relation to what is socially desirable, e.g. when per-

[20] See the paper by Bramoullé and Saint-Paul (2009) for a recent study introducing endogenous network formation model in the labor market context.
[21] The unemployment rate is $u(g) = \beta[1 - \sum_{i \in N} P_i(g)/n]$.Networks that mediate job information fluently are such that the aggregate probability $\sum_{i \in N} P_i(g)$ of finding a job through a contact is high.These networks naturally result in low unemployment (Calvó-Armengol 2004).

link cost is low, the unemployment can increase with network size. Pairwise-equilibrium networks can also be under-connected when agents hold very asymmetric positions in the network; some players may have a positive total value from the network while some player gets a negative return to it. Therefore, when link addition is welfare enhancing, it reduces unemployment. Yet, the link addition may sometimes increase unemployment and decrease welfare.

6.3. Network approach explaining phenomena in labor markets – comparison with search approach

Here, I discuss the duration dependence of unemployment and inequality in labor markets and compare the explanations of network approach and search theory to the phenomena at hand. There would be many interesting issues to consider even within labor markets, but I will restrict the analysis to this rather small area due to space constraints. Search theory is able to explain the duration dependence but not the persistent inequality. The major shortcoming of search theory is nonetheless the assumption of anonymous markets. The trends demonstrating imperfect labor markets arise because in practice, not all the employers and job searchers meet, contrary to the hypothesis of search theory. This is just what network approach accounts for. Understanding the role of networks in determining individual's position in labor markets could be one crucial factor in increasing equality in the labor markets and beyond through e.g. influencing network formation by means of transfers, for example.

6.3.1. Duration dependence of unemployment

Many of both search theoretic and network approach models recognize the fact that unemployment exhibits duration dependence and persistence. That is, the expected probability of obtaining a job decreases in the length of time that an agent has been unemployed (Calvó-Armengol and Jackson 2004)[22]. The starting points of the two approaches in clarifying this phenomenon are nonetheless very different.

The search theoretic approach explains the duration dependence of unemployment partly with the acceptance wage, the wage an individual is willing to accept and thus no longer searches for a job. Mortensen (1970) proposes that the expected duration of search, i.e. unemployment is longer the higher the acceptance wage is. This theory assumes some degree of imperfect information and that unemployed participants are nonidentical – otherwise there would be no reason to search because

[22] To illustrate the magnitude of duration dependence, Lynch (1989) finds average probabilities of finding employment on the order of 0.30 after one week of unemployment, 0.08 after eight weeks of unemployment and 0.02 after a year of unemployment (see Calvó-Armengol and Jackson 2004 for reference).

firms would offer the same wage for all openings. Here the unemployment benefit is included in the model whereas in the network approach it usually is not a decisive factor in the unemployment duration or wage differentials. The optimal acceptance wage is a function of the discount rate, the level of unemployment compensation and the proportion of jobs open to the participant, the expected duration of employment search is also a function of these same variables (Mortensen 1970). Devine and Kiefer (1991) draw a conclusion from several papers studying the relationship of unemployment benefit b and duration dependence that duration dependence is at least weakly sensitive to level of benefits. The unemployment benefit (and the taxation) should however be designed in a way that accepting a job where the wage is higher than unemployment benefit would always be optimal in money terms, which might not be the case in the low wage level. Thus the level of unemployment benefit in low-skill and low wage level can be a crucial factor in determining the duration dependence of unemployment which obviously should not occur.

Secondly, arrival rates can be important in producing variation in unemployment duration. The longer spells of joblessness experienced by some groups of workers – older, nonwhite, less educated – reflect infrequent offers since workers almost always accept an offer once an offer is received. Among Devine and Kiefer (1991) there is a direct evidence that variation in arrival rates across workers reflect variation in search effort but they cannot provide any details about how the effect operates.

In Calvó-Armengol and Jackson (2004) model where employment evolves as a function of both past employment status and the network of connections, the longer history of unemployment is more likely to come when the direct and indirect connections of an agent are unemployed and not so much from the level of acceptance wage or unemployment benefit. Thus, seeing a long spell of unemployment for some agent leads to a high conditional expectation that the agent's contacts are unemployed. This in turn leads to a lower probability of obtaining information about jobs through the social network. In fact, an agent's likelihood of being unemployed depends on her position within the network (Ioannides and Loury 2004). If there was no network connecting agents, the probability of an unemployed agent finding a job would be independent, thus simply the arrival rate a (Calvó-Armengol and Jackson 2004); then the duration dependence of unemployment would be for the most part determined by the arrival rate, the level of acceptance wage and unemployment benefit. However, compared to the network approach, the search model fails to consider some aspects that network models do. In search theory, no break-up rate is considered thus an employee can hold the job forever although on my mind the break-up rate plays a role in the determination of duration. Devine and Kiefer (1991) find that although the arrival rate might affect duration dependence, increasing number of job offers is more likely eventually to result only in higher acceptance wage, not to a lower employment rate or diminished duration dependence, but this result must hold only if the employment is at equilibrium level.

Assuming that the length of the unemployment spell can be decided by the agents and that there are always job offers available seems somewhat unrealistic. Also, the way in which workers and employers are matched is mostly left unconsidered; the matching happens more or less in a "black box" since these models do not specify where exactly the parties meet. On my mind, the theory of duration dependence in the search model is rather inadequate since it fails to consider any underlying factors affecting the duration – the characteristics of agents, the dependency of the probability of getting a job on the length of the unemployment spell etc. I believe that network approach provides a better rationalization for duration dependence of unemployment in labor markets, because network theories as such explain for example why some people might receive less job offers.

6.3.1.1. Alternative explanations

Another typical explanation for duration dependence is unobserved heterogeneity which in effect is a network externality. This means that agents have idiosyncratic features that are relevant to their attractiveness as an employee and are unobservable to the econometrician but observed by employers (Calvó-Armengol and Jackson 2004). The papers studying the relation between unobserved heterogeneity and duration dependence have obtained differing results. Consequently, whether the true explanation to duration dependence is the theory of unobserved heterogeneity or the nature of social capital, that is the quality and the number of links, is presumably not resolved (Bramoullé and Saint-Paul 2010). Bramoullé and Saint-Paul (2010) suggest a new mechanism generating true duration dependence[23].

However, the unobserved heterogeneity of agents could be in fact thought as already implicit in the network approach, because to which network(s) an agent belongs and what kind of connections s/he has tells about the intrinsic qualities of agents. In other words, if an agent is unemployed, then

[23] Duration dependence would go away once one controls for social capital in a duration model. However, this does not mean it is just another form of unobserved heterogeneity. Social capital is not an exogenous individual characteristic but its heterogeneity arises endogenously as a result of differing labor market histories across individuals. Second, the unobserved heterogeneity critique focuses on the fact that regression analysis spuriously delivers duration dependence while there is no duration dependence at the individual level: if one could estimate duration dependence with individual fixed effects it would go away. That is not true in Bramoullé and Saint-Paul model: even with individual fixed effects, duration dependence arises, because social capital is not a fixed individual characteristic but does fall endogenously over time during the unemployment spell. (Bramoullé and Saint-Paul 2010)

agents' connections are more likely to be unemployed as well and one can conclude that these agents might be of low ability, especially if the assumption of homophilous networks is taken into consideration. As a result, the network status of agent, if it is possible to reveal in studies, should uncover the unobserved heterogeneity. If such means to uncover the network status is not available the network status could in fact be one feature of unobserved heterogeneity. Also, although Calvó-Armengol and Jackson (2004) assume static network structure, the dynamic network structure might be more capable in interpreting duration dependence. As different network structures may produce different unemployment levels, obviously different structures should also produce different duration dependence properties. In other words, maybe it is the evolving network structures that create variance to duration dependence and not so much the different characteristics of agents.

Moreover, one could expect there to be differences between the dynamics of networks connecting the employed and networks connecting the unemployed (see e.g. Bramoullé and Saint-Paul 2010). The assumption e.g. of two different classes of networks is justified because the clustering properties and homophilous characteristics of "labor networks" mentioned above. The classes could differ with respect to the speed at which new links are added and old links severed and the overall degree how fast networks evolve. By assumption, the networks of the unemployed could be less dynamic because the unemployed might have less social forums where to form new links that are useful in the sense of being employed compared to the employed (coworkers, clients, partners in cooperation etc.). If the network stays unchanged, no new links between unemployed and employed people form and therefore, no job offers are received. That is, the duration dependence of unemployment might be explained by the degree of dynamics of the network structure – the more dynamic network structure the less duration dependence is experienced and the shorter is the spells of unemployment. Alternatively, if the unemployed and the employed are in the same network, the employed agents could be thought of being in the core whereas the unemployed agents in the periphery due to the different social environments. Receiving job information from the employed agents from the core to periphery is more costly and therefore also less frequent. In addition, employed agents are likely not to pass the job information so far away to the periphery but the job information is simply lost and the agents in the periphery remain unemployed. The network status of an agent could determine the length of the unemployment spell, and basically, only periphery could be shown to experience duration dependence of unemployment.

6.3.2. Inequality in terms of wage difference and dropout levels

One of the most extensively studied issues in labor economics is the persistent inequality in wages between whites and blacks. Except in duration dependence of unemployment, inequality appears in the difference of wage levels and dropout rates, for example. Even if one believes any inequality in wages between social groups can be entirely explained by differences in factors such as education, skills, and drop-out rates, one is still left to explain why those factors should differ between groups (Calvó-Armengol and Jackson 2007). Also, much richer information about the impact of networks on labor markets may be obtained by tracking wages as opposed to employment. Calvó-Armengol and Jackson (2007) find that if agents have reasonable high employment rates, then network effects will mainly be observed through their wage dynamics and correlations, as the quality of their jobs may vary dramatically even though their employment status may not. Search theoretic models mainly assume the wage difference derives from the divergence of the agents' acceptance wages.

Predicting the wage impact of the usage of networks in job search is not at all straightforward because the results on the wage effect of job search through contacts vary among demographic properties (Ioannides and Loury 2004). As an example, Korenman and Turner (1996)[24] reported that among young workers in Boston, whites who found jobs through contacts received 19 percent higher wage gains than blacks with similar characteristics. Also, Elliott (1999) found that for rather poorly educated workers, the use of informal contacts results in significantly lower wages. These outcomes can be implications from the persistent inequality of different social groups in the labor market which the profound wage differentials demonstrate. Even though the characteristics of the job seeking black and white youth in Korenman and Turner's (1996) research appear the same it is likely that the education level of the blacks' network(s) in aggregate is lower than that of the whites' network(s). This can be just one attempt among many others to explain the differences in education levels. Inequality is nevertheless permanent; the giant component structure due to preferential attachment is one rationalization why the differences between groups tend to be growing.

6.3.2.1. Correlation of characteristics in a network

Calvó-Armengol and Jackson (2007) show that improving the state of an agent's neighbors' wages or employment will improve the agent's future wages in the sense of first-order stochastic dominance. Also the wage might be increasing in the number of offers an agent has because competition between employers bids the wage up and clearly, agents with many links ought to receive more job offers as compared to poorly connected agents. Calvó-Armengol and Jackson's (2007) theorem of

[24] The two references are to be found in Ioannides and Loury (2004).

correlation pattern of wages state that any path connected agents have positively correlated wage levels in the steady state and across time, and exhibit strong association[25]. The theorem of association is a generalization of first-order stochastic dominance to random vectors. Except for wages, the positive correlation can be thought to apply also to level of education and other characteristics of agents. More accurately, an individual's position in a network is the key determinant of his or her level of activity, that is, the level of education (Calvó-Armengol et al. 2008). One rationalization why characteristics between groups of agents, social groups, should differ is the tendency of agents to form links with agents that have similar characteristics, that is, networks tend to be homophilous. Since wages and other features supposedly experience positive correlation, the wage gap or more generally the inequality between different social groups (usually different ethnic and demographic groups) can grow even higher. In addition, Montgomery (1992) states that a higher probability of a social tie and a higher percentage of educated workers decrease the probability of an individual accepting a job offer but increases wage dispersion. These factors work to perpetuate and strengthen inequality over time. Increased inbreeding by a group is shown to be associated with larger differences from other groups. Selection operates over time via network density parameters and inbreeding; individuals pass on their advantages to kin and social acquaintances (Ioannides and Loury 2004).

6.3.2.2. Wage effect of drop-out rate

Another theorem of Calvó-Armengol and Jackson (2007) which is presented in the appendix due to its complicatedness shows how persistent inequality can arise between two otherwise similar groups with different initial employment conditions. Persistent inequality is an important economicand sociological issue that search models as such fail to tackle altogether. In fact, differences in drop-out rates are an important part of the inequality in wages across races and accounting for dropouts increases e.g. the black-white wage gap (Calvó-Armengol and Jackson 2004). The positive correlation of agents' employment status also serves as a basis for understanding the differences in drop-out rates which then results in inequality in wages and employment rates. Calvó-Armengol and Jackson (2004) consider two identical networks except that one starts with each of its agents having a better employment status than their counterparts. Remaining in the labor market involves some costs and agents in the network with worse initial starting conditions have a lower expected discounted stream of future income from remaining in the network. This difference might cause some agents to drop out in the worse network but remain in the better one. Dropping out has a value 0, and agent

[25] Strong association captures the idea that better information about any of the dimensions in π leads to strictly higher expectations regarding every other dimensions in π. One implication of this is that W_i and W_j are positively correlated for any i and j in π. (Calvó-Armengol and Jackson 2007)

chooses to stay in the labor force when the discounted expected future wages exceed the costs. Due to contagion effect some agents connected to dropouts also drop out, thus a slight change in initial conditions can lead to a substantial difference in drop-out decisions and sustained differences in employment rates. Because dropping out hurts the prospects of the group further, this can have strong implications for inequality patterns between different social groups (Calvó-Armengol and Jackson 2004).

Therefore, in the network of blacks and less educated mentioned above there is probably a higher dropout rate because of the lower future discounted wage sum in comparison to the higher educated. Due to the contagion effects of the dropout decision of an individual agent dropping out results in a rather permanent wage gap between the two groups. The lower wage level accordingly discourages the blacks to educate themselves as much as would be socially optimal because the cost of education would be higher than the future wage sum given the employment and wage level. Clearly, in a networked economy, agent's actions impact other agents' decisions and possibilities because agents are interconnected; search models do not assume any game theoretic approach and therefore the only explanatory factors for the labor market outcome are essentially wage and unemployment benefit which obviously simplifies the analysis substantially. Labor markets are essentially competitive, because usually, many agents compete for the same job offer; one should at least take into consideration the decision of others when setting the level of acceptance wage.

In addition, the wage impact of using contacts in job search could be determined by the position in the network. The agents who are in the core or where the distance to the agent who has provided the job information is short the wage effect could be thought to be higher. Then again, those agents with long distance to the agent providing job information should be always expected to prefer formal channels in job search. The jobs that agents within longer distance are able to receive through contacts are probably those that more proximate job seekers have passed meaning that the job offer is less attractive in wage and other terms; which might also explain the lower wage impact. The finding of individuals forming links with people who are alike supports the fact that the black and white in fact are in different networks.

6.3.2.3. Wage effect of tie strength

The wage effect of the strength of the tie is controversial and no unambiguous results are found; ignoring the wage issue, it might be that weak ties are a more effective channel to find a job than strong ties. Granovetter (1974) argued that weak ties are superior to strong ties to providing support in getting a job because they involve a secondary ring of acquaintances that have contacts with networks outside ego's network and therefore offer new sources of information on job opportunities, weak ties thus appear to be a way for an individual to diversify her social resources. Actually, Boorman's (1975) model was the first to ask how social groups accommodate the transmission of job information, where individuals choose how to allocate effort over maintaining strong and weak ties. Under certain conditions, workers using more their weak ties than strong ties to find a job receive a higher wage (Zenou 2007). As far as the network position and strength of a tie are considered, Zenou (2007) finds that workers living far away from jobs experience higher unemployment rates than those living close to jobs because they mainly rely on their strong ties to obtain information about jobs. Workers residing closer to jobs interact more with weak ties than those residing further away from jobs. However, Montgomery (1992) suggests that weak ties increase reservation wages but it does not imply a similar relationship between wages and the type of tie used to find a job. Therefore, the (negative) correlation between the strength of the tie and the wage attained might not be due to tie strength after all but due to position in the network.

6.4. Discussion

In the general job search and matching model the fact how workers and employers learn about each other and how information is acquired is often overlooked even though a typical feature of labor markets is that there is widespread use of friends, relatives, and other acquaintances to search for jobs and it has increased over time. It seems clear that the more links an agent has to employed agents, the less likely it is that agent ends up unemployed. Nonetheless, the wage effect of being employed through contacts varies between and most probably also within a network – depending on the position in the network.

The network approach is complementary to other theories and is not meant to be exhaustive; in particular, it offers an explanation for why the workers of a particular type in a particular location (assuming networks correlate with location) might experience different employment characteristics than the same types of workers in another location, ceteris paribus. For instance, where the search theory sees duration dependence arising from the frictions in the labor market, the network approach interprets the frictions being consequence from the certain structures of networks observed in the economy. Also, the role of unobserved heterogeneity cannot perhaps be invalidated by network approach altogether though networks serve as a more multidimensional theory. These results imply that theories implementing the network approach should be developed and studied further.

The level of social interactions within a group could also explain why some individuals receive less job offers. Bayer et al. (2004) find that social interactions are stronger when individuals are more likely to interact because of education, age and the presence of children; interactions are stronger when one of the two individuals is strongly attached to the labor market, and are weaker when both are drop-outs, young or married females. Also the quality of social interaction or the quality and the nature of information can vary between networks. It might be that in the network of women, less job information is transmitted as compared to a network consisting of men, the employment level of the groups being the same. There might also be cultural differences concerning what kind of information is more likely to diffuse in a network and ultimately, to what extent the social networks are taken advantage of in job search. The cultural and cross network differences in information transmission could also explain why the wage effect of using networks can significantly vary between groups.

Differences in initial conditions combined with differences in the employment histories and with divergent network dynamics of two otherwise identical networks produce sustained inequality of wages and drop-out rates that feed each other (Ioannides and Loury 2004). That is, position in a network seems to be rather important in determining individual's professional success in terms of wage level and the decision whether to drop out or not. From the social planner's point of view, this implies that interventions in the labor market, such as providing incentives for individuals not to drop out, are likely to have noticeable and long-lasting effects in the battle against inequality. Ioannides and Loury (2004) sees that it would be more effective to target groups of agents who are highly connected, taking advantage of social attachment effects among agents. Correspondingly, institutions that seek to network otherwise isolated individuals can potentially bring about socially desirable outcomes.

7. Conclusion

In this paper, I have studied the models of network formation and networks' explicative power in rationalizing the labor market phenomena. One of the key findings here is that social structure strongly impacts the access to information, such as job information. This is the underlying explanation why networks should have an impact on the outcome of any networked markets such as labor markets that I have been discussing. The same applies to trade patterns in the networked markets of goods – the total welfare could be in line with the predictions of the neoclassical model but the surplus might be differently distributed between the buyers and the sellers. This subject is important because the traditional economic theories dominating the field assume agents act independently and are perfectly informed, and that markets are anonymous, which does not coincide with reality. Indeed, I consider the theories incorporating network approach as successful in explaining the dynamics of imperfectly functioning economy. I assume that the mismatch between supply and demand arises because all the actors in the markets do not meet. The structure of networks in the economy determines which agents operate together. Hence, the network approach is partly able to reason why there are market frictions and additionally, the explicative power of networks goes beyond than that of the market frictions.

In this concluding chapter, I shortly revise the key findings of this work. Then, I discuss the role of networks in labor markets in more detail. I conclude by deliberating the interrelationship of networks and information transmission.

7.1. The models of network formation compared

The models of network formation can be divided into two branches depending on how the links are formed. The random graph models explore the network formation on a network level and compare the emerged structure with real social networks. The basic model sees links arising independently whereas the extensions acknowledge that links form among conditional probabilities because there are underlying factors such as the tendency for people to form links with individuals having similar characteristics. Strategic form models describe the network formation process from the perspective of utility maximizing agents where the decision of forming a link essentially depends on the payoffs and costs related to it. The outcomes from the different solution concepts can be totally distinct from one another. The dilemma with the pairwise stability concept of the strategic models is that a network might be classified as stable too easily and that it only allows for one deviating action at a time. On the other hand, Nash equilibrium concept results in too many Nash stable networks of which many are irrational from the economic point of view, such as an empty network. The pairwise Nash equilibrium combining the properties of the two solution concepts could solve the problem of the concepts giving totally disjointed results, and should therefore be paid more attention to in the literature.

Although there are guidelines when random graph models and when strategic models suit modeling a certain link formation process, it might still be challenging to actually distinguish which networks form randomly and which nonrandomly. Even in strategic form models, there must be some randomness involved in that which agents actually meet before the links are formed. From the set of (randomly) met individuals an agent announces with whom s/he wants to form a link. Thus the set of agents who are in the network formation game would be determined randomly. For instance, Jackson and Wolinsky (1996) do not even assume any pregame phase but there just seem to be an (in)finite supply or some kind of line of agents in the game who meet and agree on whether to form a link or not.

7.2. Networks and labor markets

Labor markets are one typical example of networked markets where people benefit from connections with neighbors who provide information about job opportunities. Networks are a significant informal means of finding a job; hence also the focus of the labor market theories has been shifting towards network approach during the past few years. The search theoretic models provide a useful insight into how labor market should work in theory if the markets were anonymous and if only a formal job search method was available. In reality, imperfect labor markets experience duration dependence of unemployment and wage inequality which the search theoretic framework comes short of explicating. Wage differences have natural explanations such as skill differences but the distinct levels of dropping out from labor force between networks are major decisive factors in creating wage gaps between the groups. This is one example demonstrating the explicative power of the network approach. Since the labor markets actually are networked and networks play a considerable role in mediating job information the labor market models with the network approach gaining ground within the economic research is justified.

The majority of the labor market models with the network approach does not apply variable network structure but just explain these phenomena by the presence of networks. Ioannides and Loury (2004) call the variable structure as an endogenous job information network and the given structure as exogenous. Introducing variable network structure makes the analysis less straightforward and consequently, the results about the impact of contacts especially on the employment status of an agent become controversial. The key issue here is that direct contacts are beneficial because they improve an individual's information sources, but two-links away contacts are detrimental because they create competitors for the information possessed by a direct contact. The rivalry is the reason why more general analyses are difficult since the sign and intensity of the payoff spillovers are very much dependent on the geometry of the network (Ioannides and Loury 2004).

There are several mechanisms sufficient to produce segregation in labor markets. The actual skill distribution of agents is one factor producing segregation but also networks function as a source of asymmetry in labor markets. However, the inequality in the markets does not just occur passively without agents' consent. As the evidence of homophilous networks show, network formation depends on sorting on individuals' own characteristics when individuals choose whom to associate with (Ioannides and Loury 2004). This self-selection in deciding with whom an agent wishes to form a link with based e.g. on individual's own academic ability or social background most certainly increases inequality. This kind of sorting is however seen as egalitarian by most of the people, presumably because it is based on individuals' free choice of association (Ioannides and Loury 2004). Recognizing that the inequality in labor markets and generally in the society is also due to individuals' own choices and not just due to pure discrimination e.g. from the employers side makes the political intervention to the network formation more complicated and harder to justify.

Nevertheless, whether it is due to individuals' own self-selection mechanism or network dynamics one can note that many phenomena are self-enforcing since economy is organized into networks. Agents' actions shape the structure of networks but network structure also affects the way in which individuals act. The very structure of the network seems self-enforcing since the structure tends to converge into a giant component and most networks experience fat tails. Because of link formation mechanisms like triadic closure and homophilous features of networks, agents similar to each other tend to group together and even reside in the same geographical areas (see picture 6). The structure thus "enables" for the inequality to be unevenly distributed. Some agents seem to have many good quality links in terms of education and employment while other agents in minority have only few links of worse quality in terms of utility to the agent. Therefore, also dropping out is self-enforcing – every agent dropping out from labor force lowers the future discounted value of staying in the labor force of other agents. Eventually, agents with initially worse conditions are even worse off.

What the research of labor markets with the network approach should account for though is the increasing role of internet in job search. The structure of networks in labor markets should be variable because e.g. social networking sites provide a low-cost forum for links to form compared to that of physical links; thus it is very likely that the changes in the network structure are significant. In addition, it might be that internet causes dissimilar links to form than would form in a normal physical contact. That is, links crossing the borders of different social classes or educational levels might form that would otherwise not form. If this is so, social networking sites could actually diminish the length of the path between nodes so that the border between the core and the periphery somewhat blurs. My educated guess is that links formed via social networking sites must be formed more randomly than the ones formed "nonvirtually". Therefore, assuming that internet does play a significant role in the employment process, the random formation models are very relevant in the labor markets. Accordingly, the part of random graph models should be explored more in detail.

7.3. Networks and flow of information

Social networks serve as a channel for information transmission. In addition, the structure of the society matters to the economic outcome; the extent to which a society is segregated across different groups can be critical in determining things like how quickly information diffuses, and the extent to which there is under-investment in human capital, to name a few points (Curarrini et al. 2007).

Access to information is heavily influenced by social structure. For example, many social networking sites are designed so that they promote and support information diffusion from one agent to another. The information one is able to access depends on the position in the network. As a source of information networks most likely reduce the problem of information asymmetry and thus the common problem of moral hazard because one can trust the source of information, because in a network, trust is likely to emerge as Granovetter (1974) finds. Networks do not however provide for perfect information; agents are not perfectly informed contrary to the assumption of neoclassical theory but the level of information depends on the overall structure of the network and the position of the agent in the network. In fact, the information asymmetry between members of a network not connected to each other (or very distant from each other) can be even greater than in the hypothetical setting without networks. The fact that the society is organized into networks causes individuals to have asymmetric positions in the economy with respect to access to information which might result in persistent inequality between individuals.

Furthermore, the position of an agent compared to that of the other agent's position in the network could define the level of information asymmetry since agents close to each other and homophilous agents tend to have correlated information (Golub and Jackson 2008). The less distant agents the more correlated information, and also the degree to which agents are connected with the same agents could determine the level of asymmetry.

When reflecting the results of this work, it has also become clear that networks affect the behavioral patterns well beyond labor markets. Calvó-Armengol and Jackson (2004) find that the results found in the labor market model could in fact be applied to phenomena like smoking or getting involved in criminal activity. Also, Stovel and Fountain (2006) find that skill differences between groups and social segregation in the form of biased or closed networks are both associated with high levels of segregations within firms. Due to segregation within a firm, workers cannot be assumed to possess perfect information. The lack of perfect information might affect a firm's profit decreasingly which might as a consequence result in lower economic welfare. As said, networks can also determine how new products diffuse by affecting the purchase decision of individuals, particularly when the decision is affected by peers. For these reasons, social structure determines the outcome of many economic transactions.

7.4. Suggestions for future research

However, identifying the network effects is not at all simple. Many effects that one thinks are due to networks might be due to some other underlying factors. Several recent economic studies emphasize network effects as neighborhood effects; they examine whether it is appropriate to associate geographical proximity with the facilitation of information flow. Bertrand, Luttmer and Mullainathan (2000) emphasize methods that allow them to distinguish between the effects of networks from those of unobservable characteristics of individuals and of the communities where they live in their study about the impact of social networks on welfare participation. They attempt to distinguish between the effects of geographical proximity and of information transmission made possible by proximity. They find that individuals who interact more with others speaking the same language are therefore more likely to be influenced by other members of that group; they interpret these findings as the evidence of network effects. Yet, there are also studies where network effect is considered as unimportant.

All in all, it might be hard to study the impact of networks due to the overall difficulty to determine who belongs to which network and who is connected to whom. There might not be the tools available to tackle the whole network structure or the network position of an individual. In addition, what one interprets as a network effect might indeed just be due to the geographical proximity of two agents and vice versa.

For research purposes one needs more sophisticated tools than graphs to denote and explore the network structure if such do not already exist. On the other hand, revealing the structure of social online networks ought to be less complicated because of documented information about the links than that of physical networks. The structure of virtual networks could provide new insights also about the structure of physical networks. Therefore, the role of internet in the formation of networks and accordingly the social networking sites deserve more attention in the study. The inclusion of social online networks might both simplify and complicate the analysis in that social online networks and physical networks can be highly correlated, but again, due to the low costs of link formation, the structure of virtual networks might be altering much more compared to physical networks.

Indeed, even though it was not within the frame of my work, discussion about the role of institutions in contributing to the formation of networks should earn further consideration. That is, economy's social and informational infrastructure, such as internet applications facilitating or impeding agents' access to resources can have a noteworthy effect on which networks form and how quickly and through which route the information or resources flow and consequently, how much there is information asymmetry.

Additionally, as also Ioannides and Loury (2004) state, the implications of tie strength in endogenous job information networks literature should be researched more in the literature[26]. Whether the tie strength actually impacts the outcome of individuals in the labor markets is still not clear. Also, both in the labor information networks as well as in the general literature of network formation it would be interesting to see some hybrid models incorporating features from both random and strategic form models. All in all, the network approach should be implemented into various economic settings.

[26] In fact, Tassier and Menczer (2008) discuss about the effects of social network structure on inequality; they study how randomness of the network and also how randomness of the flow of job information affects the labor market outcome and consequently the emerging inequality.

References

Robert J. Aumann and J.H. Dreze (1974):" Cooperative Games with Coalition Structures", International Journal of Game Theory 3.

Robert J. Aumann and Roger B. Myerson (1988):" Endogenous Formation of Links between Players and of Coalitions: an Application of the Shapley Value", The Shapley value: essays in honor of Lloyd S. Shapley, edited by Alvin E. Roth, Cambridge University Press

Venkatesh Bala and Sanjeev Goyal (2000):" A Noncooperative Model of Network Formation", Econometrica, Vol. 68, No. 5

Albert-Lászlo Barabási and Réka Albert (1999): "Emergence of Scaling in Random Networks", Science, Vol. 286

Patrick Bayer, Stephen Ross and Giorgio Topa (2004):" Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes", Working Paper, Yale University.

Julian Besag (1974):" Spatial Interaction and the Statistical Analysis of Lattice Systems", Journal of the Royal Statistical Society, Series B 36.

Bernheim, B., Bezalel Peleg, and Michael Whinston (1987):"Coalition-Proof Nash Equilibria I: Concepts," Journal of Economic Theory, Vol. 42, 1-12.

Marianne Bertrand, Erzo Luttmer and Sendhil Mullainathan (2000):" Network Effects and Welfare Culture", Quarterly Journal of Economics, CXV, August.

Francis Bloch and Bhaskar Dutta (2009):" Communication Networks with Endogenous Link Strength", Games and Economic Behavior, Vol. 66

Francis Bloch and Matthew O. Jackson (2006):" Definitions of equilibrium in Network Formation Games", International Journal of Game Theory, vol. 34, no. 305

Scott Boorman (1975):" A Combinatorial Optimization Model for Transmission of Job Information through Contact Networks", Bell Journal of Economics and Management Science, Vol. 6.

Yann Bramoullé and Gilles Saint-Paul (2010):" Social Networks and Labor Market Transitions", Labour Economics, Vol.17.

Antoni Calvó-Armengol (2004):" Job Contact Networks", Journal of Economic Theory, Vol. 155, No.1

Antoni Calvó-Armengol and Matthew O. Jackson (2004):" The Effects of Social Network on Employment and Inequality", The American Economic Review, Vol. 94, No.3

Antoni Calvó-Armengol and Matthew O. Jackson (2007):" Networks in Labor Markets: Wage and Employment Dynamics and Inequality," Journal of Economic Theory, Elsevier, Vol. 132, No. 1

Antoni Calvó-Armengol and Yves Zenou (2005):" Job Matching, Social Network and Word-of-Mouth Communication," Journal of Urban Economics, Elsevier, vol. 57, No.3.

Sergio Currarini, Matthew O. Jackson and Paolo Pin (2007):" An Economic Model of Friendship: Homophily, Minorities and Segregation", Working Papers, Department of Economics Ca' Foscari University of Venice, No. 20/WP

Theresa J. Devine and Nicholas M. Kiefer (1991): "Empirical Labor Economics: The Search Approach", Oxford University Press.

Bhaskar Dutta and Matthew O. Jackson (2001):” On the Formation of Networks and Groups”, Draft forthcoming in the Models of the Strategic Formation of Networks and Groups, Heidelberg: Springer Verlag.

Bhaskar Dutta, Anne van den Nouweland and Stef Tijs (1998):” Link Formation in Cooperative Situations”, International Journal of Game Theory, vol. 27, 245-256

David Easley and Jon Kleinberg (2009): “Networks, Crowds and Markets – Reasoning about a Highly Connected World, Draft version of September 29, 2009, to be published in the Cambridge University Press, Spring 2010

Erdös and Rényi (1959): “On Random Graphs”, Publicationes Mathematicae Debrecen, Vol. 6, 290-297

Frank, O., and D. Strauss (1986):” Markov Graphs”, Journal of the American Statistical Association, Vol. 81.

Andrea Galeotti and Sanjeev Goyal (2002):“Network Formation with Heterogeneous Players”, Tinbergen Institute Discussion Paper, TI 2002-069/1

Benjamin Golub and Matthew Jackson (2009):” How Homophily Affects Learning and Diffusion in Networks”, Fondazione Eni Enrico Mattei Working Paper, Sustainable Development Series, No. 35.

Sanjeev Goyal, Marco van der Leij and José Luis Moraga-González (2006):” Economics: An Emerging Small World”, Journal of Political Economy, vol. 114, no. 2

Mark Granovetter (1973):” The Strength of Weak Ties”, The American Journal of Sociology, Vol. 78, No. 6

Mark Granovetter (2005):” The Impact of Social Structure on Economic Outcomes”, Journal of Economic Perspectives, vol. 19, no. 1

Hendricks, K., Piccione, M., and Tan, G. (1995):” The Economics of Hubs: The Case of Monopoly. Review of Economic Studies. Vol. 83-99.

Yannis M. Ioannides and Linda Datcher Loury (2004):” Job Information Networks, Neighborhood Effects and Inequality”, Forthcoming in the Journal of Economic Literature.

Matthew O. Jackson (2008a): ”Social and Economic Networks”, Princeton University Press.

Matthew O. Jackson (2008b): "Network formation", The New Palgrave Dictionary of Economics, Second Edition, Palgrave Macmillan. The New Palgrave Dictionary of Economics Online. 08 June 2009
<http://www.dictionaryofeconomics.com/article?id=pde2008_N000135>
doi:10.1057/9780230226203.1174

Matthew O. Jackson (2003):” Allocation Rules for Network Games”, Working Papers 1160, California Institute of Technology, Division of the Humanities and Social Sciences.

Matthew Jackson and Asher Wolinsky (1996): ”A Strategic Model of Social and Economic Networks”, Journal of Economic Theory 71.

Alan Kirman(1997):” The Economy as an Evolving Network”, Journal of Evolutionary Economics, Vol.7, 339-353.

Alan Kirman and A. Vignes (1991):” Price Dispersion. Theoretical Considerations and Empirical Evidence from the Marseille Fish Market”, Issues in Contemporary Economics, editor K.J. Arrow, Macmillan, London.

Rachel E. Kranton and Deborah F. Minehart (2001): ”A Theory of Buyer-Seller Networks”, The American Economic Review Vol. 91, No. 3

Michael König, Claudio Tessone and Yves Zenou (2009):” A Dynamic Model of Network Formation with Strategic Interactions”, CEPR Discussion Papers, No. 7521.

Thomas A. Lubik (2009):” Estimating a Search and Matching Model of the Aggregate Labor Market”, Economic Quarterly, Vol. 95, No. 2.

Marco A. Marini (2007): ”An Overview of Coalitions and Networks Formation Models for Economic Applications”, WP-EMS Working Papers Series in Economics Mathematics and Statistics.

James D. Montgomery (1994):” Weak Ties, Employment and Inequality: An Equilibrium Analysis”, American Journal of Sociology, Vol. 99, No.5.

James D. Montgomery (1991):” Social Networks and Labor-Market Outcomes: Toward and Economic Analysis”, The American Economic Review, Vol. 81, No. 5.

Roger B. Myerson (1977):”Graphs and Cooperation in Games”, Mathematics of Operations Research, Vol.2, No.3

Roger B. Myerson (1991):”Game Theory: Analysis of Conflict”, Harvard University Press

John Nash (1951):” Non-Cooperative Games”, The Annals of Mathematics, Second Series, Vol. 54, No. 2

John Nash (1953):” Two-Person Cooperative Games”, Econometrica, Econometric Society, Vol. 21, No.1.

von Neumann and Morgenstern (1944):” Theory of Games and Economic Behavior”, Princeton University Press

M. E. J. Newman (2002):” Random Graphs as Models of Networks”, arXiv: cond-mat/0202208v1, working paper

M.E.J. Newman, D.J. Watts, S.H. Strogatz (2002): ”Random Graph Models of Social Networks” Proceedings of the National Academy of Sciences of the United States of America, Vol. 99, No. 1.

Qin, C. (1996): “Endogenous Formation of Cooperation Structures,” Journal of Economic Theory, 69, 218-226.

Garry Robins, Pip Pattison, Yuval Kalish and Dean Lusher (2007):” An Introduction to Exponential Random Graph (P*) Models for Social Networks”, Social Networks, Vol. 29.

Brian W. Rogers (2005):” A Strategic Theory of Network Status”, Mimeo. MEDS, Northwestern University.

Marco Slikker, Anne van den Nouweland (2001):” Social and Economic Networks in Cooperative Game Theory”, Springer.

Tom Snijders, Philippa Pattison, Garry Robins and Mark Handcock (2006):” New Spefications for Exponential Random Graph Models”, Sociological Methodology , 99-153

Katherine Stovel and Christine Fountain (2006):” Hearing about a Job: Employer Preferences, Networks and Labor Market Segregation”, Working Paper, Allacademic.

Ning Sun, Walter Trockel, Zaifu Yang (2008):” Competitive Outcomes and Endogenous Coalition Formation in an *n*-Person Game “, Journal of Mathematical Economics, Vol.44, No. 7-8.

Troy Tassier and Filippo Menczer (2008):” Social Network Structure, Segregation, and Equality in a Labor Market with Referral Hiring”, Journal of Economic Behavior & Organization, Vol. 66, No. 3-4.

Anne van den Nouweland (2004): ”The Formation of Communication Networks in Cooperative Games”, Group Formation in Economics – Networks, Clubs and Coalitions, Edited by Gabrielle Demange and Myrna Wooders, Cambridge University Press 2005.

Duncan J. Watts and Steven H. Strogatz (1998):”Collective Dynamics of ‘Small-World’ Networks”, Nature 393.

Zenou (2007):” Social Interactions and Labor Market Outcomes in the Cities”, CEPR Discussion Papers 6129.

Ezra W. Zuckerman (2003): “On Networks and Markets by Rauch and Gasella, eds.”, Journal of Economic Literature, Vol. XLI, June.

Web:

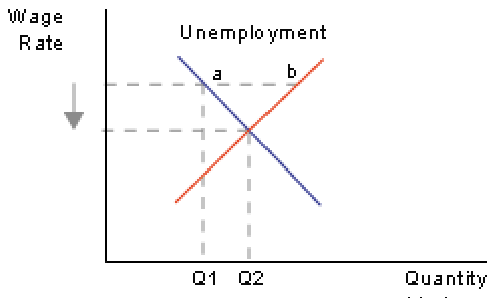
The New Palgrave Dictionary of Economics
http://www.dictionaryofeconomics.com/dictionary

http://tori.tekes.fi/blogs/attachment/482/8783/3142/Sosiaalinen_media_uudistaa_liiketoimintaa_-tiedote.pdf
(14.10.2009)

http://mathworld.wolfram.com/MarkovChain.html/
www.math.ucsd.edu/~fan/algo/sigma.html

Appendix

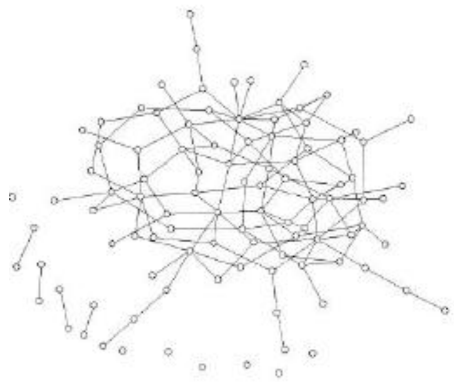
Labor market equilibrium in neoclassical theory



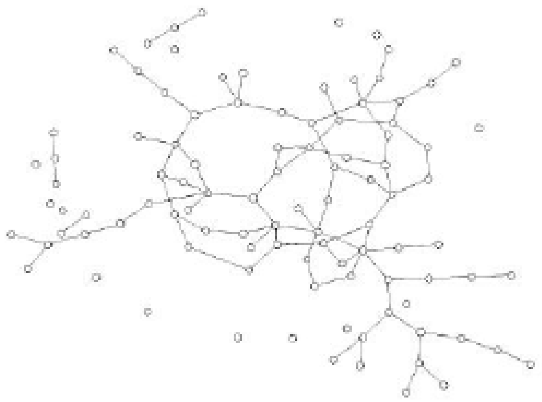
Picture 8. Wages are initially too high and there is unemployment of ab. This causes wage rates to fall and employment increases as a result from Q1 to Q2. Any unemployment left in the economy would be purely voluntary unemployment. (New Palgrave Dictionary of Economics)

Different graphs

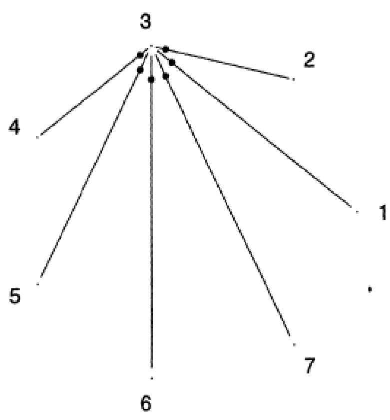
Here, when comparing the two graphs, picture 8 and picture 9, one can notice that the average degree of a vertex in a Bernoulli random graph is seemingly lower than in the left hand small world graph. That is, the small world graph clearly exhibits clustering and fat tails.



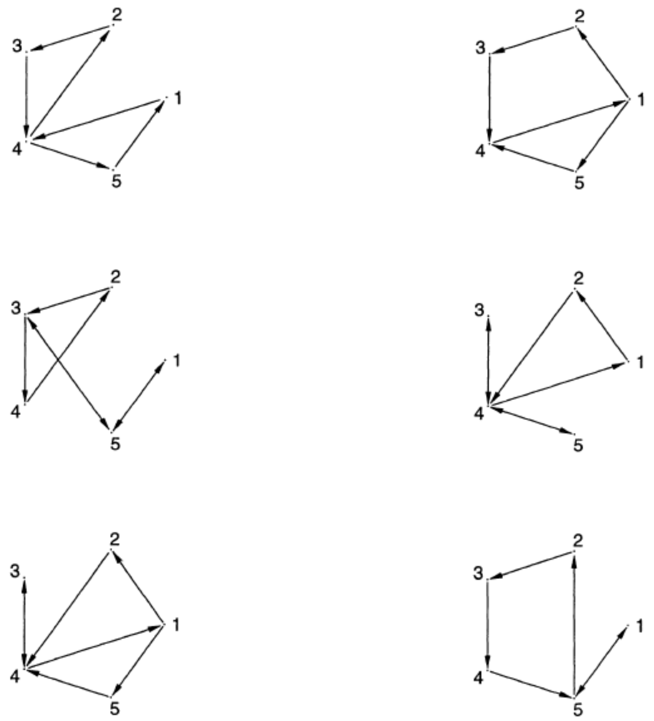
Picture 9. **Small world graph.**



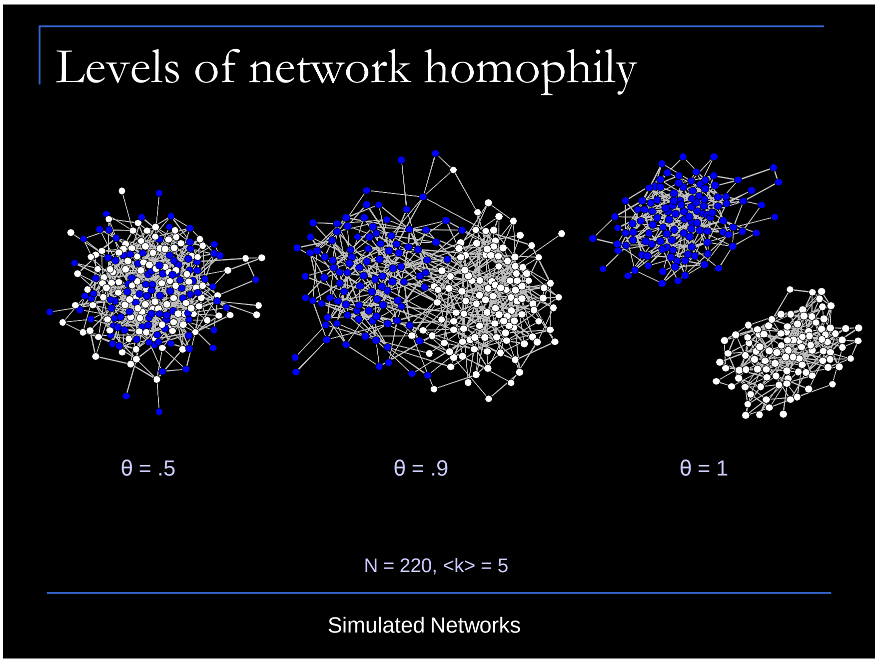
Picture 10. **Bernoulli random graph.**



Picture 11. **Center-sponsored star. (Bala and Goyal 2000)**



Picture 12. **Examples of Nash networks in the linear payoff case.** The number of Nash networks increase quite rapidly with n. For example, there are 5, 58, 1069 and in excess of 20 000 Nash networks as n takes values 3,4,5 and 6, respectively. (Bala and Goyal 2000).



Picture 13. **Levels of network homophily (Stovel and Fountain, University of Washington 2006)**

Calvó-Armengol and Jackson (2007) model

W_{it} is the wage of the agent i at time t. The vector $w_t = (w_{1t}, \dots, w_{nt})$ is a realization of the wage levels at t. The random variable S_{it} is the employment status of an agent at time t. The vector $s_t \in \{0,1\}^n$ describes the employment status at time t.

$b_i \in (0,1)$ is the breakup rate. $p_{ij}(W_{t-1})$ is the probability that *i* originally hears about a job in a given period and then it is eventually *j* that ends up with an offer for that job. $p_i(w) = \sum_j p_{ji}(w)$ is the expected number of offers that i will get when the wage state in the last period is w.
O_{it} denotes the new opportunities that *i* has in hand at the end of the hiring process in period t. The wage of agent i evolves as

$$W_{it} = w_i(W_{i,t-1}, O_{it})S_{it}$$

Wage patterns and dynamics
To measure the wage patterns and dynamics, Calvó-Armengol and Jackson (2007) use the measure of association. A probability measure μ describing a random vector is *associated* if $Cov_\mu(f, g) \geq 0$.
Starting from the steady state distribution, there is a strictly positive correlation between the wage statuses of any path connected agents and at any times, for large enough T. That is, for any times t and t' and large enough T,

$$Cov^T[W_{it}, W_{jt'}] > 0$$

where i and j are path connected and Cov^T is the covariance associated with the T-period subdivision starting at time 0 under the steady state distribution μ^T.

Dropping out and long run inequality
Let $d_i \in \{0, 1\}$ denote *i*'s decision of whether to stay in the labor market. Each agent discounts future wages at a rate $0 < \delta_i < 1$ and pays an expected discounted cost $c_i \geq 0$ to stay in. Agents dropping out get a payoff of zero. An augmented economy is *(N, p, b, c, δ)*, where *c* and *δ* are vectors of costs and discount rates.
When an agent *i* exits the labor force, we reset the *p*'s so that $p_{ij}(w) = p_{ji}(w) = 0$ for all *j* and *w*, but do not alter the other p_{kj}'s. The agent who drops out has his or her wage set to zero. 13
Therefore, when an agent drops out, it is as if the agent disappeared from the economy.
Fix an augmented economy *(N, p, b, c, δ)* and a starting state $W_0 = w$. A vector of decisions *d* is an *equilibrium* if for each $i \in \{1, \ldots, n\}$, $d_i = 1$ implies

$$E\left[\sum_t \delta_i^t W_{it} \,|\, W_0 = w, d_{-i}\right] \geqslant c_i$$

If $d^*(w)_i = d^*(w')_i = 1$, then the distributions of i's wages and employment W_{it} and S_{it} for any t under the maximal equilibrium following w' first-order stochastic dominate those under the maximal equilibrium following w, with strict dominance for large enough t if $d^*(w)_j \neq d^*(w')_j$ *for any j who is path connected to i. In fact, for any increasing* $f: \mathbb{R}_+^n \to \mathbb{R}$ *and any t*

$$E^T\left[f(W_t)\,\middle|\,W_0 = w', d^*(w')\right] \geqslant E^T\left[f(W_t)\,\middle|\,W_0 = w, d^*(w)\right],$$

with strict inequality for some specifications of c and δ.

This shows how persistent inequality can arise between two otherwise similar groups with different initial employment conditions.